ENVIRONMENTAL RESTORATION PROJECT PEER REVIEW PROCESS ASSESSMENT

THESIS

Paul A. Schantz, Captain, USAF

AFIT/GEE/ENV/96D-18

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

DTIC QUALITY INSPECTED 3

Wright-Patterson Air Force Base, Ohio

AFIT/GEE/ENV/96D-18

ENVIRONMENTAL RESTORATION PROJECT PEER REVIEW PROCESS ASSESSMENT

THESIS

Paul A. Schantz, Captain, USAF

AFIT/GEE/ENV/96D-18

19970317 031

Approved for public release, distribution unlimited

DTIC QUALITY INSPECTED 3

The contents of this thesis are technically accurate, and no sensitive items, detrimental ideas, or deleterious information are contained therein. Furthermore, the views expressed in this thesis are those of the author and do not reflect the official policy nor views of the US Air Force, Department of Defense, nor the US Government.

AFIT/GEE/ENV/96D-18

ENVIRONMENTAL RESTORATION PROJECT PEER REVIEW PROCESS ASSESSMENT

THESIS

Paul A. Schantz, B.S., M.S., Captain, USAF

Presented to the Faculty of the Graduate School of Engineering
of the Air Force Institute of Technology
In Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Engineering and Environmental Management

WILLIAM B. NIXON, Major, USAF, Ph.D.
Committee Member

ROBERT P. STEEL, Ph.D.
Committee Member

STEVEN T. LOFGREN, Lt Col, USAF, Ph.D.
Committee Chairman

AFIT/GEE/ENV/96D-18

ENVIRONMENTAL RESTORATION PROJECT PEER REVIEW PROCESS ASSESSMENT

THESIS

Presented to the Faculty of the Graduate School of Engineering
of the Air Force Institute of Technology
Air Education and Training Command
In Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Engineering and Environmental Management

Paul A. Schantz, B.S., M.S.
Captain, USAF

December 1996

Approved for public release, distribution unlimited

Acknowledgments

This research project was an enjoyable learning experience due to the support and flexibility I was given by my advisor, Lieutenant Colonel Steve Lofgren. He was a constant source of motivation. He provided very timely guidance without limiting the learning process. Further, his assistance in developing the structure of this research was very crucial to the final product being clear and readable.

I was also very fortunate to have Major Brent Nixon and Dr. Robert Steel as involved committee members. Major Nixon provided helpful advice and assistance throughout the research process. Dr. Steel's help with research design and statistical methods was invaluable.

My sponsors from the Air Force Center for Environmental Excellence were especially instrumental in the beginning of this research effort in narrowing my research focus and choosing a meaningful area to pursue. I especially thank Major Dan Welch and Dr. Stan Hewins for their time and efforts. Dr. Bill Sweet and Mrs. Elizabeth Hewins were also helpful in shaping the early stages of this research. Thank you to Richard Trevino, from the Air Education and Training Command, and Wayne Ratliff, from the Air Force Materiel Command, for their time and support.

Finally, I thank my wonderful wife, Laura, whose constant support and understanding was critical to making the thesis process an enjoyable success. And no acknowledgment is complete without thanking God for this gracious opportunity and for the gift of living in freedom in the United States of America!

Paul Schantz

Table of Contents

Page

Acknowledgments ii

List of Figures vii

List of Tables viii

List of Acronyms ix

Abstract xi

I Introduction 1-1

Background 1-1
Policy 1-2
Peer Review Requirement 1-4
Problem Statement 1-6
Research Objectives 1-8
Research Questions 1-8
Significance 1-9
Limitations 1-11

II Literature Review 2-1

Peer Review History 2-1
  Manuscripts 2-2
  Grants 2-6
  Construction Projects 2-6
Peer Review Applications 2-7
  Manuscripts 2-7
    Manuscript Review Evaluation 2-9
    Manuscript Review Recommendations 2-11
  Grants 2-15
  Accountancy 2-16
  Research Programs 2-17
  Construction Projects 2-20
    Categories 2-22
      Project Peer Review 2-24
        Project Design Peer Review 2-25
Peer Review Process 2-28
  Innovation versus Risk Abatement 2-28

Peer Reviewer Characteristics 2-30
Reviewer Evaluation Instrument 2-32
Summary 2-34

III Survey Development 3-1

Construct Development 3-1
Effectiveness Construct 3-1
Construct Discovery Method 3-4
Literature Review Content Analysis 3-6
Observations 3-7
Interviews 3-8
Open-Ended Feedback Forms 3-9
Data Triangulation 3-10
Seven Predictor Constructs 3-11
Focus 3-11
Agenda and Facilitator Constructs 3-14
Written Preparation 3-17
Oral Presentation 3-18
Peer Review Team 3-20
Peer Reviewer 3-22
Seven Predictor Constructs Grounded in Air Force Policy 3-23
Final Phase of Survey Development 3-25
Survey Cover Letter and Instructions 3-25
Pretest Instrument 3-26

IV Method 4-1

Research Design 4-1
Assumptions 4-1
Sample and Setting 4-2
Survey Variables 4-4
Peer Review Process Effectiveness 4-4
Peer Review Process Focus 4-5
Agenda 4-5
Facilitator 4-5
Written Preparation 4-5
Oral Presentation 4-6
Peer Review Team Composition 4-6
Peer Reviewer Characteristics 4-6
Survey Data Plan 4-7
Collection 4-7
Reduction 4-7
Follow-up 4-8

Statistical Analysis 4-8
  Multi-Item Variables 4-8
  Descriptive and Inferential Statistics 4-9
    Descriptive Statistics 4-9
    Frequency Distribution 4-9
    Test for Multicollinearity 4-10
    Test for Internal Consistency 4-10
    Linear Regression Model 4-11
    Confidence Level 4-12
    Uniqueness Index 4-12

V Results 5-1

General Survey Responses 5-1
  Air Force Center for Environmental Excellence 5-1
  Air Education and Training Command 5-1
  Air Force Materiel Command 5-2
Demographics 5-2
Peer Review Effectiveness 5-3
Descriptive Statistics and Correlations 5-5
Composite Variables Correlations 5-6
General Linear Regression Model Results 5-8
  Air Force Center for Environmental Excellence 5-9
  Air Education and Training Command 5-9
  Air Force Materiel Command 5-11
Open-Ended Items 5-12
  Air Force Center for Environmental Excellence 5-12
  Air Education and Training Command 5-12
  Air Force Materiel Command 5-13
Observation Results 5-13
  Air Force Center for Environmental Excellence 5-14
  Air Force Materiel Command 5-14

VI Discussion 6-1

Introduction 6-1
Effective Peer Review Process 6-1
General Model 6-2
  Oral Presentation 6-2
  Focus 6-2
Significant Variables by Major Command 6-3
  Oral Presentation 6-3
  Reviewer Characteristics 6-4
  Focus 6-4

  Written Preparation 6-4
  Facilitator 6-5
Peer Review Process Recommendations 6-6
  Scope and Objectives 6-6
  Peer Reviewer Mix 6-6
  Trained Facilitator 6-7
  Prepublished Agenda 6-8
  Visual and Verbal Overview 6-8
  Prioritization of Peer Review Goals 6-8
  Peer Review Team as Technical Advisory Board 6-9
  Peer Review Protocol 6-10
Future Research 6-10

Appendix A: Air Force Peer Review Process Descriptions and Focuses A-1

Appendix B: Air Force Base Conversion Agency Peer Review Observations B-1

Appendix C: Peer Review Interviews C-1

Appendix D: Air Force Base Conversion Agency Peer Review Feedback Forms D-1

Appendix E: Questionnaire E-1

Appendix F: Survey Development Procedure F-1

Appendix G: Survey Administration Approval Requirements G-1

Appendix H: Statistical Background H-1

Appendix I: Air Force Base Conversion Agency Peer Review Survey Comments I-1

Appendix J: Air Education & Training Command Survey Comments J-1

Appendix K: Air Force Materiel Command Survey Comments K-1

Appendix L: Air Force Materiel Command Peer Review Observations L-1

Appendix M: Survey Item Frequency Distributions M-1

Bibliography Bib-1

Vita Vit-1

List of Figures

Figure Page

1. Peer Review Placement in the IRP Process 1-5
2. Simplified Hierarchy of Restoration Guidelines Affecting Peer Review 1-6
3. Peer Review History 2-1
4. Important Issues 2-22
5. Design Peer Review Categories 2-23
6. Traits of a Good Reviewer 2-32
7. Air Force Environmental Program Objectives 3-2
8. Air Force Environmental Program Criteria 3-3
9. Four-Fold Peer Review Program Purpose 3-3
10. Iterative Content Analysis Steps 3-6
11. Data Triangulation 3-11
12. Populations and Samples 4-2
13. Survey Respondents' Assigned Organizations 5-2
14. Survey Respondents' Job Positions 5-3
15. Overall Peer Review Effectiveness Responses 5-4
16. Utilization of Resources Responses 5-5
17. Peer Review Process Recommendations 6-11
18. Interview Protocol C-1
19. Steps in Survey Instrument Design F-1
20. Guidelines for Survey Question Development F-2
21. Six Techniques to Increase Survey Response Rate F-3

List of Tables

Table Page

1. How Peer Review Purpose Meets Nine Criteria 3-4
2. Criterion Measure Item-Total Correlations & Scale Alpha 5-6
3. Predictor Variable Item-Total Correlations & Scale Alphas 5-7
4. Correlation Matrix 5-8
5. Results of Stepwise Linear Regression (General) 5-10
6. Results of Stepwise Linear Regression (AFCEE) 5-10
7. Results of Stepwise Linear Regression (AETC) 5-10
8. Results of Stepwise Linear Regression (AFMC) 5-10

List of Acronyms

ACEC	American Consulting Engineers Council
AETC	Air Education and Training Command
AF	Air Force
AFBCA	Air Force Base Conversion Agency
AFCEE	Air Force Center For Environmental Excellence
AFI	Air Force Instruction
AFIRM	Air Force Installation Restoration Management Meeting
AFIT	Air Force Institute of Technology
AFMC	Air Force Materiel Command
AFPD	Air Force Policy Directive
ASCE	American Society of Civil Engineers
ASFE	Associated Soils and Foundation Engineers
CASE	Coalition of American Structural Engineers
CERCLA	Comprehensive Environmental Response, Compensation and Liability Act
CEVR	Civil Engineering Environmental Restoration
CPA	Certified Public Accountant
DERA	Defense Environmental Restoration Account
DERP	Defense Environmental Restoration Program
DoAF	Department of the Air Force
DoE	Department of Energy

EO	Executive Order
EPA	Environmental Protection Agency
FS	Feasibility Study
IRP	Installation Restoration Program
MajCom	Major Command
OMB	Office of Management and Budget
RA	Remedial Action
RCRA	Resource Conservation and Recovery Act
USAF	United States Air Force

Abstract

Installation Restoration Program (IRP) projects often cost over $250 thousand, and projects over $1.5 million are common. To ensure these projects are risk based, technically sound, and cost effective, the Air Force instituted a peer review program in 1992. The objective of this research is to describe and analyze the peer review process.

Through triangulation of data from interviews, observations, official and academic documents, and surveys, seven constructs were discovered: focus, agenda, facilitator, written preparation, oral presentation, team characteristics, and reviewer characteristics.

A questionnaire was used to gather perceptions of peer review effectiveness--the criterion variable--and of the seven constructs. A total of 141 surveys administered with a 50% response rate (N = 70).

Linear regression is next used to assess the predictiveness of the seven constructs. Oral preparation and focus predict 68% of the variance in peer review effectiveness; no other constructs were significant. Variation in the peer review process between major commands is explored.

Last, recommendations are made to improve the current peer review process.

ENVIRONMENTAL RESTORATION PROJECT PEER REVIEW PROCESS ASSESSMENT

Chapter 1

Introduction

Background

In the early 1990s, the Headquarters of the United States Air Force Environmental Restoration Branch (HQ USAF/CEVR) noted a trend of questionable spending on Installation Restoration Program projects. This, coupled with recent cuts in the Defense Environmental Restoration Account (DERA), indicated a need for review of Installation Restoration Program projects. In 1992, HQ USAF/CEVR instituted a peer review program to ensure Installation Restoration Program projects were both cost-effective and environmentally-appropriate; the peer review program was delegated to the major commands for administration (Furlong: 1995).

The major commands each developed peer review programs to oversee environmental projects (Appendix A). The peer review programs vary widely between major commands from the Air Force Materiel Command, which conducts in-house peer reviews, to Air Combat Command, which conducts peer reviews utilizing a contractor. These peer review programs have developed in different directions without an established conduit for cross-flow of information.

With the large number of base closures in the 1990s, the Air Force Base Conversion Agency was tasked with oversight of a wide range of environmental restoration projects. To ensure fiscal responsibility and technical feasibility, the Air Force Base Conversion Agency instituted a peer review program separate from the major commands that was implemented by the Air Force Center for Environmental Excellence (AFCEE). AFCEE implemented a peer review program that utilized reviewers from AFCEE, the Environmental Protection Agency (EPA), contract, and other external sources.

In 1995, AFCEE was appointed as the Air Force Office of Primary Responsibility to establish guidelines and best business practices for the Air Force peer review program (AFCEE, 1996:1). To meet this responsibility, AFCEE developed and distributed the Peer Review Guidance Document in January, 1996. The guidance document establishes: the purpose of the peer review program; the roles and responsibilities of Air Force organizations to implement the peer review program; and guidance for implementing an environmental project peer review.

Policy

The major pieces of legislation that generate the current environmental restoration requirement are: the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA) and the Resource Conservation and Recovery Act (RCRA). In addition, two executive orders (EOs)--12088, Federal Compliance With Pollution Control

Standards and 12580, Superfund Implementation--generate current restoration requirements. The Superfund Amendments and Reauthorization Act (SARA) revises and extends CERCLA to establish the Defense Environmental Restoration Program (DERP). The DERP's funding mechanism is the Defense Environmental Restoration Account (DERA) (Department of the Air Force [DoAF], 1993b:3-7).

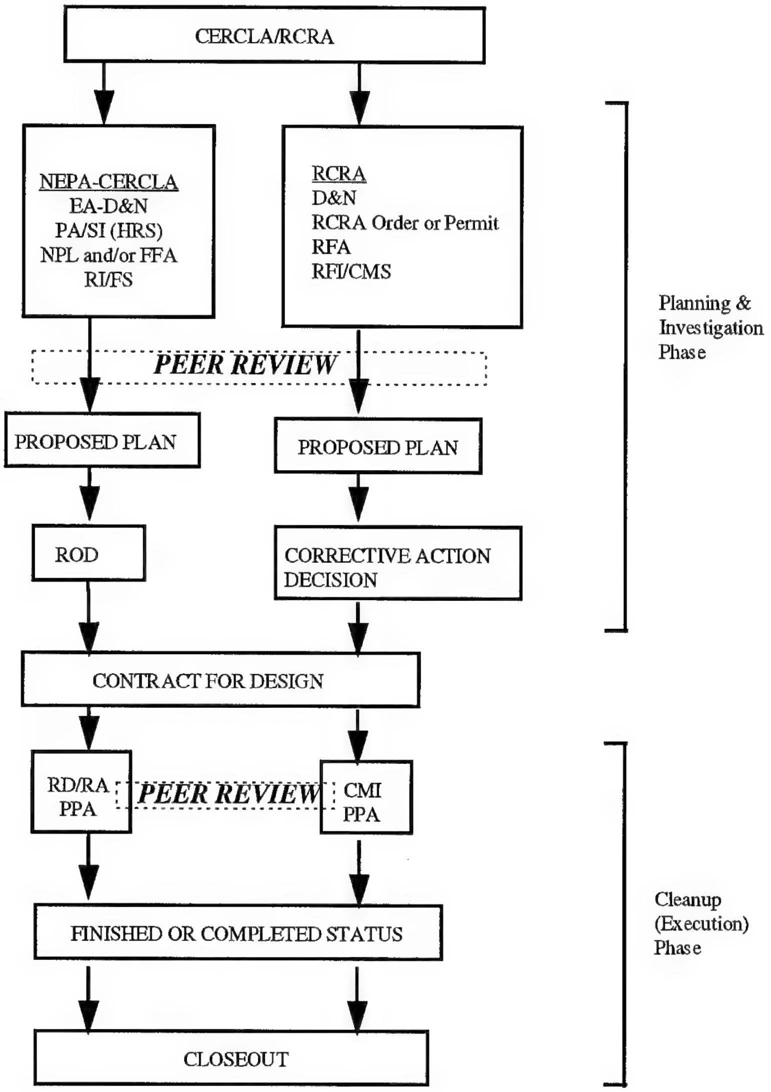
The Installation Restoration Program (IRP) was established by the Air Force to meet the requirements of CERCLA, SARA, RCRA, EO12088, and EO12580. The purpose of the Installation Restoration Program is to "identify, evaluate, investigate, and clean up hazardous substances or wastes used at sites in the past. The purpose is to protect public health and the environment" (DoAF, 1993b:3-6). The Air Force has established "sites finished" as its measure of merit (process output metric) for the Installation Restoration Program process. Sites finished are either closed-out or in steady-state long term operations and maintenance such as pump-and-treat (DoAF, 1993b:3-5).

The U.S. Air Force Installation Restoration Program Remedial Project Manager's Handbook (DoAF, 1993b:3-15), described later as the Handbook, states that regulators may pursue remediation of hazardous sites under either CERCLA or RCRA and that overlap is possible. The Resource Conservation and Recovery Act (RCRA) of 1976 "normally applies to currently active practices involving solid and hazardous waste management. RCRA may be applied by regulatory agencies to require remediation for past improper hazardous waste disposal practices and spills that resulted in a threat to the environment and human health" (DoAF, 1993b:3-15). To mitigate lost work if sites can fall under either RCRA or CERCLA, the Handbook (DoAF, 1993b) recommends a proactive approach to designing a remediation plan that meets the requirements of both CERCLA and RCRA.

Peer Review Requirement

Figure 1 shows the Installation Restoration Program project phases and pathways to close-out. While it is outside the scope of this thesis to describe in detail the steps in the Installation Restoration Program close-out process, this figure is included to show the points in the Installation Restoration Program process at which peer review is recommended by AFI 32-7020, *The Environmental Restoration Program.*

Peer review is recommended by the Handbook (DoAF, 1993b) at two locations in the Installation Restoration Program close-out process. First, peer review should be instituted "toward the end of the Feasibility Study (FS) to review the alternatives considered and select the best cleanup alternative, and second, after the Remedial Design (RD) to review the scope and cost for Remedial Action (RA)" (DoAF, 1994c:4). AFI 32-7020 is the Air Force Instruction that implements the Installation Restoration Program directed by Air Force Policy Directive (AFPD) 32-70, *Environmental Quality.* Figure 2 shows a simplified hierarchy of the policy implementing the Air Force's Installation Restoration Program. Peer review programs must satisfy the requirements of each of these policy documents.


CERCLA/RCRA
NEPA-CERCLA
EA-D&N
PA/SI (HRS)
NPL and/or FFA
RI/FS
RCRA
D&N
RCRA Order or Permit
RFA
RFI/CMS
Planning & Investigation Phase
PEER REVIEW
PROPOSED PLAN
PROPOSED PLAN
ROD
CORRECTIVE ACTION DECISION
CONTRACT FOR DESIGN
RD/RA PPA
PEER REVIEW
CMI PPA
Cleanup (Execution) Phase
FINISHED OR COMPLETED STATUS
CLOSEOUT

Abbreviations

CMI	Corrective Measures Investigation
CMS	Corrective Measures Study
D&N	Discovery & Notification
EA	Environmental Assessment
FFA	Federal Facilities Agreement
HRS	Hazardous Ranking System
NPL	National Priorities List
PA/SI	Preliminary Assessment/Site Investigation
PPA	Past Project Activities
RD/RA	Remedial Design/Remedial Action
RFA	RCRA Facilities Assessment
RFI	RCRA Facilities Investigation
ROD	Record of Decision

Figure 1. Peer Review Placement in the IRP Process
(Adapted from: DoAF, 1993b:3-8)

Document	Description
CERCLA and RCRA & EOs	Federal Acts & Executive Orders
DERP	DoD Program
IRP	Air Force Program
AFPD 32-70	Air Force Policy
AFI 32-7020	Air Force Instruction

Figure 2. Simplified Hierarchy of Documents Affecting Peer Review

Problem Statement

The major commands and the Air Force Base Conversion Agency (AFBCA) have developed separate peer review programs with little collaboration. Improvements can be made through the sharing of information and best business practices. To this end, AFCEE was given the responsibility to develop guidance for the Air Force's peer review program. The Air Force would benefit from an effort to develop metrics to continuously improve the peer review process. Currently, the Air Force Base Conversion Agency is using an output metric of dollars saved as an indication of peer review program effectiveness. Dollars saved is calculated by subtracting the post-review cost estimate from the original funding request document.

The "dollars saved" output metric does not foster continuous improvement of the peer review process because it fails to meet the attributes as required by the Metrics Handbook (DoAF, 1993b:2-1), such as timeliness and driving the appropriate action. Generally speaking, what you measure is what you get. Using an output measure that focuses on cost savings will inappropriately focus the peer review on spending less money versus getting the most performance per dollar spent.

Output metrics are not practical for several reasons. First, the projects reviewed are usually in the planning or design stage. Second, while projects are required to be estimated using such tools as the Remedial Action Cost Engineering and Requirements System (RACER), cost estimates are historically inaccurate. Third, additional performance specifications may be added by the peer reviewers to meet the requirements of the project. Due to these reasons, comparing the cost estimate before the peer review to the cost estimate after the review does not necessarily indicate improvement. An increase in price may indicate a range of causes such as: an improved cost estimate, additional performance required to meet project requirements, or estimate variance.

Due to downsizing across the Air Force, all organizations must do "more with less." To focus resources and improvement in the peer review process, the Air Force Center for Environmental Excellence has been appointed the lead Air Force organization for improving the peer review process. The Air Force Center for Environmental Excellence developed the Peer Review Guidance Document (1996) with a goal to foster continuous improvement of the peer review program. To help meet this end, the current research seeks to determine the peer review process constructs that are the most important for improvement. These constructs would provide a process metric to foster continuous improvement of the peer review process. "The assumption is: if process performance is monitored and improved, the quality of the products and services will improve" (DoAF, 1991:3-3).

Research Objectives

The first objective of this research is to review the literature for lessons learned from related fields to begin building the project peer review process body of knowledge. This includes discovering the constructs that are important to an effective peer review process. Since academic literature on peer review programs is scarce, the objective is focused more on discovery than on applying established theory. The second objective is to review Air Force policy and define the criterion variable, peer review process effectiveness. The third objective is to determine statistically which constructs are the strongest predictors of the perceived effectiveness of the peer review process. The underlying objective of this research is to provide the Air Force, and others utilizing project peer review, a tool to narrow improvement focus to maximize the utilization of limited resources. The academic objective of this research is to suggest constructs that are important in predicting peer review effectiveness and to rank the constructs by their predictive ability.

Research Questions

Specifically, the objectives of this research are to learn:

- What lessons can be gleaned from peer review outside the Air Force?
- What characteristics are deemed important by Air Force peer review experts and through observation of the Air Force environmental peer review process?
- What constructs capture these lessons learned, expert opinion, and observations?
- What are the parameters of these constructs?

- How is the environmental peer review process effectiveness defined by Air Force policy?
- Which constructs most strongly predict the effectiveness of the Air Force environmental peer review process?

Significance

Since the Air Force Center for Environmental Excellence is currently responsible for the peer review program, improvements in peer review guidance will impact all major commands and the Air Force Base Conversion Agency. The current output metric used by the Air Force Base Conversion Agency does not foster timely continuous improvement. Recent interviews with the people responsible for the peer review programs at the various major commands have indicated people are eager to find new ways to improve the peer review process. This research will have broad application since the environmental projects under peer review are very similar. Further, several major commands have peer review processes very similar to the process used by the Air Force Base Conversion Agency (implemented by the Air Force Center for Environmental Excellence), the Air Education and Training Command, and the Air Force Materiel Command.

The tightening environmental budget is further reason to find new ways to stretch the dollar. Recent programs presented at the Air Force Installation Restoration Management Meeting (AFIRM) in 1996 stressed leveraging resources to meet the heightened environmental requirements with decreased dollars. Peer review was presented as one tool to ensure fiscal responsibility, technical feasibility, and risk-based decision-making.

To put the magnitude of potential savings in proportion, the FY1996 Air Force portion of the Defense Environmental Restoration Account budget was $365 million (Air Force Materiel Command, 1996:26). Installation Restoration Program projects often cost over $250,000, and projects over $1.5 million are common (Air Force Materiel Command, 1996:54). With these expensive projects, the USAF--as a public steward--has the responsibility to provide fiscal, technological and managerial oversight to Installation Restoration Program projects. Peer review provides this oversight through an expert panel of reviewers. By refining and improving the peer review process, the Installation Restoration Program can be improved which helps make tight dollars go farther. In order to improve the peer review process, the Air Force needs a metric to foster process understanding and to motivate action to continuously improve the process. This research is the first step in developing this metric.

The literature review indicates academic study on project design peer review is deficient and even more so on environmental project design peer review. This research begins to fill a void in the body of knowledge. Manuscript peer review for scientific and medical journals is probably the most researched area of peer review. Research on manuscript peer review was accelerated by the First International Congress on Peer Review in 1989. Approximately 30% of the papers presented in this first congress were based on opinion. At the Second International Congress on Peer Review in 1993, there was a trend toward more scientific investigation in the papers presented (Smith, 1994:144).

Project peer review has very little academic study with the exception of the debate initiated by Zallen (1990) with his "Proposal for Structural Design Peer Review." However, this academic article is qualitative in orientation. There are few rigorous scientific experiments, or quantitative studies, in the literature. Project peer review is a recent development, with the Associated Soils and Foundation Engineers (ASFE) initiating a peer review program to reduce liability and improve quality of construction in 1978 (American Consulting Engineers Council, 1988:1). The peer review most similar in objective and scope to environmental project peer review is project design peer review. The guidelines for project design peer reviews were published in 1990 by the American Consulting Engineers Council in cooperation with the American Society of Civil Engineers. These guidelines are empirically based upon the experience of consultants working in the field.

Limitations

This research evaluates the characteristics of the current peer review processes used by the Air Force Center for Environmental Excellence, the Air Education and Training Command, and the Air Force Materiel Command during Spring and Summer 1996. Peer review processes are variable both in time and across organizations. Readers must ascertain for themselves whether the results of this study can be generalized based on the similarities between the evaluated peer review programs and their own.

Chapter 2

Literature Review

Peer Review History

The history of formal peer review dates back to the 18th century with a gradual, disjointed refinement to the formal peer review used by many organizations today. This section traces the implementation of peer review through medical and scientific manuscript review, grant review, project review, and finally culminates with the Air Force's adoption of peer review for the environmental process. Figure 3 illustrates the development of peer review.

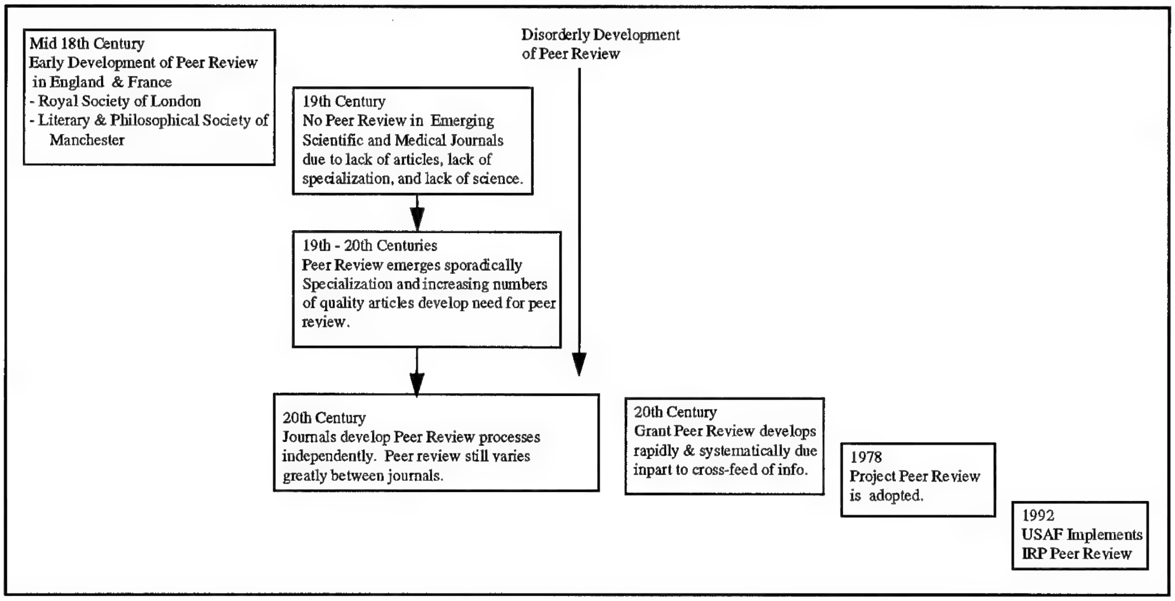

Mid 18th Century
Early Development of Peer Review in England & France
- Royal Society of London
- Literary & Philosophical Society of Manchester
Disorderly Development of Peer Review
19th Century
No Peer Review in Emerging Scientific and Medical Journals due to lack of articles, lack of specialization, and lack of science.
19th - 20th Centuries
Peer Review emerges sporadically Specialization and increasing numbers of quality articles develop need for peer review.
20th Century
Journals develop Peer Review processes independently. Peer review still varies greatly between journals.
20th Century
Grant Peer Review develops rapidly & systematically due inpart to cross-feed of info.
1978
Project Peer Review is adopted.
1992
USAF Implements IRP Peer Review

Figure 3. Peer Review History

Manuscripts. Two principle authors of the history of peer review, Kronick and Burnham, differ in their perspective on the history of peer review. Kronick describes a formal peer review process used by scientific journals in the 1700s that used standardized procedures to choose which manuscripts to publish. In contrast, Burnham gives an account of journals in the United States in the 1800s that used an informal "seat-of-the-pants" review to select reports for publishing (Knoll, 1990:1330). While these two authors offer different perspectives of peer review history, their accounts are not contradictory. Peer review did not develop systematically like many other aspects of science. Peer review has developed, for the most part, without cross-feed of information. Each journal developed and shaped their individual peer review process at different rates and though different patterns of development.

Generally, peer review has existed since people began to communicate new ideas which had to pass the scrutiny of others (Kronick, 1990:1321). New ideas are first peer reviewed by contemporaries in the field who either oppose or accept the new theory. If consensus is achieved, the new idea is integrated into the body of knowledge; if not, the new idea is either dropped, further tested, or modified and resubmitted for review. While this process is effective for small amounts of new information, as information grows and becomes more specialized, peer review must also become more specialized and systematic.

Peer review effectively started in the 18th century, emerging out of a need for a review of the emerging scientific hypothesis in Europe. In 1752, the Royal Society of London introduced the concept of peer reviewing scientific manuscripts prior to publication

(Kronick, 1990:1321). Even in its infancy, the importance of structure and guidelines in the peer review process was clear. The society established their "Committee on Papers," to review all articles published in the Philosophical Transactions. The committee was composed of five members; other experts could be consulted as required. The Literary and Philosophical Society of Manchester also initiated a peer review policy in the 18th century that was similarly formal in peer review policy (Kronick, 1990:1321).

Although the peer review system introduced by the Royal Society of London is similar to the system used by most journals today, it was not systematically adopted by emerging scientific and medical journals. Peer review policies came into different scientific journals and practices over an extended period of time and were not standardized until the 20th century (Burnham, 1990:1327). According to Burnham (1990:1324), during the 19th century medical journals emerged concurrently with the peak of personal journalism in general magazines and newspapers. The early medical journals adopted many of the practices of the popular newspapers such as editors using the journals to express their personal opinions.

The medical journals grew to such large numbers that "John Shaw Billings, in 1879, thought that it was 'as useless to advise a man not to start a new journal as it is to advise him not to commit suicide.'" (Burnham, 1990:1324). The early emphasis was on quantity not quality of articles; however toward the end of the 19th century, as newspapers began to focus on news and moved away from opinion, the medical journals followed suit and began to more closely scrutinize submitted articles.

The resistance to peer review resulted from: a lack of articles; medical editors opinion that medical doctors did not require review; the "right to publish;" and the editor's desire to use the journal as an educational instrument. In the early days of medical journals, editors had to seek articles as opposed to the current requirement to sift through many good articles to find the best to publish (Burnham, 1990:1325). The medical journals were very resistant to peer review. The editors of these journals insisted that anyone possessing a medical degree had no reason to defer to a colleague by virtue of the degree (Burnham, 1990:1325).

Resistance to peer review is found in any area being peer reviewed. There is always a resistance to being reviewed by ones "equals." Furthermore, editors of medical journals often served as the salesmen for organizations and were often pressured into printing the member's articles, regardless of quality. This pressure was constant in the discord between organizations who felt they had a "right" to be published and the editors who felt they had a duty to uphold quality.

Another factor leading to closer scrutinization of submitted articles was the specialization of science. As science developed through the end of the 19th century and into the 20th century, it became increasingly specialized, resulting in editors and their staffs being unable to effectively review articles. The pace of technological and scientific development made it impractical for a staff to be current in the wide range of areas needing review.

The transition from editorial review to editorial staff review to external peer review was not systematic nor expedient. Even as late as the first half of the 20th century, Smith

Ely Jelliffe, editor of the Journal of Nervous and Mental Disease, would not send journal submissions out for peer review even though he lacked authority in the subfields of neurology and psychiatry (Burnham, 1990:1324). Many editors had a barrier of pride that prevented them from utilizing external peer review. These cutting-edge scientific developments often involved lives of real people. Insufficient review of scientific ideas often put lives at risk which was immoral (Burnham, 1990:1324).

In the 20th century, increased specialization and the abundance of submitted articles pressured many journals into using a peer review process. In the early 20th century, the Journal of the American Medical Association evolved from seeking any manuscripts to accepting only the best through the use of peer reviews (Burnham, 1990:1325). Peer reviews became increasingly prevalent through the 20th century as science continued to develop at an accelerated pace. "In the physical and biological sciences, for example, a 1962 survey of 156 journals in 13 countries found that 71% made some use of referees" (Zuckerman and Merton, 1971:75). This was a great shift from the days when editors had to search for articles to fill their journals. A second factor was the increased specialization of professions. No longer could a general medical practitioner effectively review articles from specialized medical and science fields. The following case study demonstrates how the need for the correct mix of specialists on a peer review committee developed:

> In 1901, in a case which expertise, not overcrowding of pages, was the issue, Harold Moyer, editor of Medicine, told of an article that appeared in an eastern journal in which the author claimed to have discovered a substance, ureine, that caused uremia. Moyer reported, three different investigators took time out to verify the falsity of the claim; the foolish author had confused precipitating chemical solids with urine solids. Clearly the model of the outside expert was

acceptable to at least this one editor--Moyer--when he recognized that a chemist might have prevented the problem (Burnham, 1990:1326).

Grants. The history of grant peer review is in sharp contrast with the disorderly development of journal peer review. Grant peer review evolved as communities of scientists collaborated to develop a uniform process. Grant peer review was introduced in the United States through the work of the Carnegie Institute of Washington, DC in the 20th century (Burnham, 1990:1328). Grant peer review is currently often mandated in legislation. A recent example of this was a requirement for peer review of all grant requests from a fiscal year 1993 U.S. Army breast cancer initiative (Schwartz et al, 1995:558).

Construction Projects. In the later part of the 20th century, as construction projects became increasingly specialized and immense, the need for engineering design peer review developed. In 1971, Boston, Massachusetts instituted a system of peer review for structural designs exceeding $1 million (Zallen, 1990:208; Barnes, 1995:1085; Johnson, 1995). This peer review requirement was later adopted by the state of Massachusetts. Following the lead of Massachusetts and other states, project peer review was formalized by the Association of Soil and Foundation Engineering in 1978 to meet the growing needs of the specialized civil engineering and construction field (Preziosi, 1988:46). From 1978 through 1987, Connecticut bore witness to several major structural failures which led to a requirement for special inspections and project peer review starting in 1988 (Carlson, 1995:1102). Other states similarly adopted formal peer review requirements for large engineering projects. Unlike the development of journal peer review programs, Army project design peer reviews have a formal procedure established for transferring information between programs (Department of the Army, 1994:2-5).

Peer Review Applications

The preceding discussion of peer review history primarily revolved around developments in manuscript review. The manuscript peer review process was first historically and is currently the most studied application of peer review. Many of the characteristics of manuscript peer review are general and can be uses in other applications. These characteristics include whether the system is series or parallel; whether those under peer review are blinded; and the means of measuring successful peer review processes.

Five peer review applications were studied: manuscript, grant, accountancy, research programs, and construction projects. Grant and accountancy peer reviews are minimally applicable to AF environmental peer review and were accordingly minimized in this review. Research program peer review is more applicable but is not thoroughly covered in the literature. Construction project peer review is directly applicable to environmental project peer review and is accordingly emphasized. Project design peer reviews are the most similar to environmental project peer reviews.

Manuscripts. According to Weller (1990:1346), who studied 159 journals, the fundamental process of editorial peer review is similar among indexed medical journals. First manuscripts are sent outside to an accepted list of reviewer(s). The reviewers read the manuscripts and make a recommendation to accept, reject, or revise, based on their expertise and forward the results to the editorial staff. The final decision is made by the editor or the editorial staff based on the review results (Weller, 1990:1346).

Journal editors use two primary systems: series and parallel. In the series peer review system, the editor sends the manuscript to a single referee for review. If the referee recommends the article, the editor almost always publishes it. If the reviewers reject the article, the editor almost always requests revisions before sending the manuscript back to the same or to a new reviewer for a second opinion. In the parallel referee system, the editor sends the manuscript to two or more reviewers simultaneously. If their assessments agree, the editor usually follows their recommendations. If their assessments disagree, the editor usually sends the manuscript to another reviewer to reach consensus or asks for further revision from the author (Hargens, 1990:1348-49).

There are three types of reviewing processes used with respect to revealing the author's and reviewer's identity. In a nonblind system, both the reviewer and the author are known to each other; in this system the referee signs the review, which is returned to the author. In the double-blind system, neither the reviewer, nor the author are aware of the other's identity; the manuscript is sent to the referee with the name blanked, and the author is not informed who reviewed his manuscript. In the single-blind system, the author does not know who reviewed his manuscript (Blank, 1991:1041).

Blank studied the effectiveness of double-blind versus single-blind manuscript reviewing and found that reviewers are more critical and less likely to recommend acceptance of the manuscript for publishing when they are unaware of the author's identity (Blank, 1991:1041). McNutt et al (1990:1370) also found blinding increases the effectiveness of peer review. While closer scrutiny is gained from the double-blind system, it is only useful for non-verbal exchange of information and it creates a lack of trust between the reviewer and the party under review.

Manuscript Review Evaluation. Smith has noted that there is little high quality peer review research (Smith, 1994:144). Drummond Rennie, deputy editor at The Journal of the American Medical Association, advanced research in peer review by setting-up the First International Congress on Peer Review in 1989. In the first congress, about 30% of the 35 papers selected for presentation were opinions, not the results of investigations (Smith, 1994:144). JAMA sponsored the Second International Congress on Peer Review in September 1993 (Rennie, 1994:91). The second congress attracted 110 papers with a heavier focus on scientific investigation (Smith, 1994:144). The Third Congress on Peer Review is scheduled for September 1997 (Rennie, 1995:987). In regards to the science behind peer review, Knoll notes that editorial peer review is more of an art, than a science; it is a different process each time and for each place. Further, there is no precise means to select the best method for peer review for all situations (Knoll, 1990:1330).

Knoll states that one of the greatest problems with trying to improve or study peer review is the lack of available outcome measures. While process measures are available to assist in continuous improvement, it is valuable to have outcome measures to verify the improvement (Knoll, 1990:1330). A further complication is in defining a good outcome of a peer review. Should the top priority be minimizing rater variability, or should it be to minimize errors in review, or a combination of the two?

Many peer review researchers debate which metric is the best process measurement for continuous improvement. By analyzing a database of 301 pairs of reviewer ratings from 120 different manuscripts, Gallagher (1989:iv) found that in a parallel referee system the overall recommendation by the referees to accept or reject was in agreement 79% of the time. Gallagher argues that this is an indicator of the consistency and validity of the peer review process. Atkinson (1994:148) counters that the consistency of reviewer's recommendations as an overall statistic is irrelevant; what is critical is how much innovation is stifled and whether progress in cutting-edge technology is delayed.

Luce (1993) suggests the effectiveness of peer review should be measured by measuring Type I and II errors instead of measuring rater reliability. Type I errors (α) are the proportion of articles that are filtered out by peer review when in-fact they are later recognized as correct. Type II errors (β) are the proportion of articles that are not filtered out by peer review and are published and are later found to be deeply flawed (Luce, 1993:399). While this metric sounds effective, it is difficult to apply in many situations because it is an outcome measure. It may be years before a theory that was filtered by the peer review process is demonstrated as a success; likewise, it may take years to realize that a published theory is invalid.

Rejection rate is an important characteristic of the peer review filter to optimize the reviewing process, to minimize α and β. In the electronic industry, filters are designed to minimize unwanted noise, while maximizing "let through" and minimizing delay

(Atkinson, 1994:148). Likewise, peer review should have these goals: to minimize β (published flawed articles), while maximizing the proportion of "non-flawed" articles ($1-\alpha$), and minimizing delay.

Manuscript Review Recommendations. Research into peer review has led to a wide range of recommendations for the manuscript peer review process ranging from Sinclair's argument (1993), that the peer review process should be discarded from journals and publish all, to Atkinson's development of the argument that the peer review process should be tightened due to the occurrence of fraud. With an increased competition for a shrinking research budget, fraud has become increasingly prevalent. "The recommended procedures to combat fraud are typically authoritarian: referees should be more expert, more vigilant, more censorious, etc." (Atkinson, 1994:147).

This reaction, to tighten the peer review filter, focuses on correcting the symptom, with little investigation given to the overall peer review process. To utilize a system view, first peer review process goals should be determined. Then, using these goals as guidance, the peer review process should be analyzed for the corrections that will lead us toward these ends. The goals of the peer review process with respect to filtering are to maximize the number of non-erroneous articles published and to minimize the number of articles published with significant errors. Atkinson (1994) warns against over zealous rejection of new work.

> Other forms of bias and misrepresentation which effect science communication and therefore science are probably more prevalent than fraud and certainly more harmful. Unreasonable opposition from referees which causes suppression of new ideas is no better than fraud: indeed it is worse, because a suppressed idea is lost whereas fraudulent claims can be discounted (Atkinson, 1994:147-8).

According to Atkinson (1994), limiting fraud should not be a top priority in the peer review process.

Cicchetti (1991) looks at the reliability of peer review for manuscript and grant submissions across disciplines. His article opens active discussion of the importance of the reliability of peer review and he suggests using standard sets of acceptable scientific criteria to improve inter-referee agreement and the peer review process. When a high level of inter-referee agreement is desirable, Cicchetti (1991), Daniel (1993:67), and Oxman et al (1991:95) suggest using standard sets of acceptable scientific criteria (Cicchetti, 1993:403). Oxman et al (1991:95) further recommends reviewer training to improve consistency of peer reviews.

Hirschman agrees that there should be criteria for reviewers, but feels that the criteria should be customized for each review.

> It is especially important that care is taken in the design of evaluation criteria for peer review. These criteria must be explicit, fully defined, and communicated effectively to all reviewers so that the reviewers' ideas about effectiveness can, as much as possible, be brought in line with those of the sponsoring agency (Hirschman, 1994:86).

This difference between Cicchetti's and Hirschman's recommendations for criteria is probably due to the fact that Cicchetti was recommending the criteria for a much narrower field of peer review than Hirschman.

Luce (1993:399) argues against standard reviewer criteria because they will limit the effectiveness of the review. Luce states that editors attempt to maximize individual differences between reviewers to reduce the likelihood that grossly erroneous findings are published. Luce notes that the more divergent the raters, the more likely the review will detect errors. Further, some raters are better than others at detecting certain types of errors. This "is the reason experienced editors (e.g., Bailar, Kiesler, and Roediger) select referees with different special abilities--to catch different types of errors" (Luce, 1993:399).

Luce's main point is that focus should not be on increasing rater consistency and reliability to improve the peer review process since divergent reviewers are desirable and lead to a better peer review filter. Given that the primary goals of the referee system are to both maximize worthy articles getting published and minimize unworthy articles getting published (Chubin and Hackett, 1990:44)--coupled with the fact that editors minimize these errors by selecting reviewers that differ substantially--leads automatically, and desirably, to low reliabilities (Luce, 1993:400). For peer review to be effective, it must reduce both Type I and Type II errors. This will ensure that the most error-free research gets funded or promoted (Hirschman, 1994:86).

Sinclair (1993:400) argues that peer review censorship should be discarded completely. He modifies Cicchetti's work and concludes that peer review should be terminated and all articles should be published. Since the reliability of review increases as a direct function of the number of reviewers, the maximum reliability is reached with the maximum number of reviewers--i.e., all interested readers (Sinclair, 1993:400). Sinclair proposes that journals terminate the peer review filter and publish all scientific articles. Prior to publication, the manuscript might still be sent out for review. If the article was severely critiqued and rejected, the author would still have the opportunity to revise the manuscript and publish the article alongside the reviewer's critique. The larger body of readers would have the opportunity to decide for themselves the validity of the research. Sinclair goes further to say that rejecting articles is similar to throwing-out data and misleading the scientific community (Sinclair, 1993:400).

Cicchetti's response (1993:401) to Luce and Sinclair gives a different perspective to reliability and the importance of the peer review process. Cicchetti suggests that Luce's argument that reliability in peer review is not expected nor desired is not substantiated since the major scientific journals rely on reviewer reliability as a metric for continuous improvement. Granted, it is important that reviewers are selected for least likely consensus, but when consensus is reached, the editor has an even stronger recommendation since consensus was not "predetermined" through the selection of similar reviewers (Cicchetti, 1993:403).

Editors using both the parallel and the series referee systems depend on agreement between raters to force the publication decision. For the parallel referee system, the editor usually goes with the consensus; if the reviews are in conflict, the editor usually sends the manuscript to another reviewer. Similarly, in the series system, if the first referee recommends rejection of the manuscript, the editor typically sends the manuscript to a second reviewer. If the second reviewer is in consensus with the first reviewer, then the editor usually rejects the article. If the first and second reviewer give different recommendations, then the manuscript is usually sent to a third reviewer and the editor usually goes with the majority (Cicchetti, 1993:403).

Garfunkel et al (1990) conducted an experiment to study the effectiveness of manuscript review. He had a similar pair of reviewers analyze manuscripts that had already been approved through the peer review process. He found that 80% of the manuscripts were accepted the second time; however, the initial peer review failed to address important issues uncovered by the second review. One possible solution to this problem, recommended by Garfunkel, is to increase the number of reviewers.

Atkinson (1994) recommends against anonymity. He argues that the practice of accomplishing peer review anonymously has negative ramifications on the party under review's trust in the review system. Further, if the reviewer is truly a "peer," the identity will surface sooner or later. If the reviewer is not a "peer," then the anonymity is dangerous in the sense that it masks potential ignorance on the part of the reviewer (Atkinson, 1994:155).

Peer review should not be used for more than it is designed for: minimizing type I and type II errors. "There are fundamental weaknesses in the notion of peer review: peer status is illusory, perception is fickle, and impartiality is unnatural. Peer opinion is not, *ipso facto*, scientific argument" (Atkinson, 1994:157-8). To foster cooperation and build trust in the peer review process, the burden of proof in peer reviewing should rest with the reviewer (Atkinson, 1994:158). Further, Type I error (incorrect filtering of valid "non-flawed" articles) should be minimized.

Grants. Harold Varmus, National Institute of Health director, is implementing both a "triage" approach to sorting grant applications and a "just-in-time" rule for providing data. The triage approach involves a first-pass quick review that eliminates 30 to 50% of the applications off the top before they are sent to the panel for complete review (Marshall, 1994:467). This cuts down on time spent by peer review panels and decreases the turn-around time for rejected grant applications. The just-in-time rule for providing data alters the submission process so that just those grant applications that make it through the first-pass of the triage are required to submit detailed funding and planning data (Marshall, 1994:467). This cuts down on the time wasted by grant applicants who are rejected in the first-pass--time that can be spent doing research or refining the grant application. The National Institutes of Health (1980) uses a process similar to the "triage" approach for contracts in emerging scientific fields.

Accountancy. Peer review is used in the accounting professions as both a means of quality control and as a means of increasing the accounting firm's credibility with the customer. The peer review determines whether the accounting "firm's system of quality control for its accounting and auditing practice met the objectives of the quality control standards established by the American Institute of Certified Public Accountants (AICPA)" ("Official Releases: Peer Review Standards Interpretations," 1995:118). To put the magnitude of the number of these peer reviews in perspective, over 10,000 CPA firms are scheduled for quality or peer review every year through the AICPA (McCabe, 1991:111).

The magnitude of this program requires a systematic approach to administering peer review. Through the use of standards and guidelines, the peer review process can efficiently review the firms with a cost savings. Further, a systematic program for monitoring accountant practices helps the American Institute of Certified Public Accountants maintain high quality performance (Official Releases: Peer Review Standards Interpretations, 1995:133; Peer Review Board of the American Institute of Certified Public Accountants, 1995:3103). Another benefit of systematic peer review can be an increase in efficiencies for CPAs from exchanges of ideas and exposure to different types and sizes of firms. Both the peer reviewer and the reviewer's firm benefit from this information flow (Marshall, 1994:111).

The process for accountancy peer review is different. As noted in "Official Releases: Peer Review Standards and Interpretations" (1995:118), the CPA (Certified Public Accountant) peer reviewer is required to:

> test administrative and personnel files; review selected engagements, including the relevant working files and reports; interview firm personnel; access other confidential matter, as appropriate; and communicate his or her conclusion to senior members of the reviewed firm at an exit conference.

Research Programs. While manuscript, grant, and accountancy peer reviews offer insights into the peer review process, investigation of program and project peer review will hopefully result in ideas that can be directly transferred to environmental peer review. Cozzens (1987:72) forms a matrix to describe the interrelationships between the different requirements for peer review at the programmatic level. The customer or stakeholder may be inside the organization owning the project or external to it. An internal customer would be within the organization responsible for the project being peer reviewed, while an external customer would be outside the organization responsible for the project. Further, the project may be specific or general in scope. A specific project is applied research where there is a specific goal of the program. A general scope applies to basic research where only a general area of science is specified.

Cozzens (1987) conducts a narrow survey of major firms and uses their respective peer review processes as case studies to investigate how the customer, whether internal or external, in combination with project scope, effect the type of evaluation procedures. Cozzens notes that reviews with external customers are generally formal and utilize external experts, while reviews with internal customers are informal and utilize in-house experts. This retrospective case study found that the reviews were qualitative with the exception of external reviews of general projects/programs which were more quantitative.

Cozzens (1987:73) case study of IBM uncovered attributes of their integrated peer review process that may be useful in other organizations. First, IBM integrates peer review into the corporate management in budgeting and goal setting. They are concerned with "developing the right technologies at the right time in the future" (Cozzens, 1987:73). Second, each project receives continuous monitoring and evaluation by line managers. Third, IBM has instituted "IBM fellows" who are senior scientists given the freedom to move throughout the company, participate in a range of activities, and report to upper management. Further, IBM uses external experts to assist with its research programs on an as-needed basis (Cozzens, 1987:73).

Hirschman investigates the peer review process at the programmatic level of organizations. Peer review programs have become the institution that balances the rapid advancement of science on the one hand with accountability of public funds and safety on the other. (Hirschman, 1994:80; Chubin and Hackett, 1990:46). Organizations use peer review for both internal and external purposes. Peer review can be used both for allocation of internal resources or for legitimization of organization's projects or programs to external customers.

Hirschman observes through case studies that peer review is often an afterthought. "It is important to realize that peer review is not just an accessory that can be plugged into an organization. The process must be understood and integrated into a broader understanding of the need for evaluation within organizations" (Hirschman, 1994:80). Hirschman finds that the use of peer review results by decision-makers should be stated explicitly and monitored to ensure the results are used effectively (Hirschman, 1994:82). It is crucial that decision-makers have the flexibility to override peer review findings if other factors influence the decision; however, it is equally important that decision-makers do not abuse this flexibility and disregard the peer review results based on a whim.

The European Economic Community (EEC) uses a five-member peer review team to evaluate R&D subprograms. The ECC Solar Energy R&D review panel uses a three tier review process. On the first tier, they evaluate all sectors on the basis of status reports. On the second tier, they conduct a more in-depth review of a sample of the contracts based on more detailed information. On the third tier, the panel conducts a complete evaluation on a very small sample of contracts by analyzing all pertinent information. Unfortunately Cozzens (1987:75-76) does not evaluate the effectiveness of this process nor was it found elsewhere in the literature. It is included in this analysis because of its uniqueness and possible application to other project peer reviews.

Construction Projects. Project peer review is defined as a "structured, comprehensive and thorough fact-finding process conducted by one or more senior professionals who are separate and independent from the organization preparing a project design" by the American Society of Civil Engineers' manual of professional practice, Quality in the Constructed Project (Preziosi, 1988:46). It is important to communicate the purpose and scope of the project review to the reviewer(s).

Like Gallager (1989), who states that peer reviewers should be more jurors than detectives, Zallen (1990:212) argues that an engineering peer review is "not intended as a comprehensive design check; it is only intended to verify that the design is conceptually correct and that there are no major errors. The peer reviewer assumes that the details of the design are correct, the peer review is to verify the overall design concept. However, Zallen (1990:213) recommends reviewers perform independent calculations of a representative sample of the project details. This gives the reviewer both a fresh perspective and allows the reviewer to use the method most familiar to him. If the results disagree, then the reviewer checks the original calculations to find the discrepancy.

The General Services Administration (GSA) recently established peer review for major public projects. Edward Feiner, GSA's chief architect, structured two design reviews to test two different peer review processes. The first was a two-stage design review conducted by a panel of three experts. The first stage of review was conducted at the preliminary-schematic stage of design. The second stage of review was held once the design was final. The second design relied on an individual independent design professional. Feiner found that both review processes were effective, but noted that the synergy of ideas in the group proved to be the most effective forum (Nesmith, 1993:29). Zallen (1990) also recommends a two-stage approach similar to that used by Feiner.

Preziosi illuminates several pitfalls to avoid in engineering peer review. One is the "'that's not the way I would have done it' syndrome" (Kline in Preziosi, 1988:47). This occurs when an expert reviewer is overly critical of work due to competition or difference in method. This syndrome can be lessened through peer reviewer education and through a careful selection of peer reviewers. Another problem illuminated by Preziosi is sensitivity to criticism; especially when project peer review involves interaction between the reviewer(s) and the reviewed engineer. "One such problem, according to Hillis, is that a review suggests incompetence. 'With the hint of incompetence, the engineer goes into a defensive posture when dealing with the reviewer. Then there ends up being a communication problem'" (Preziosi, 1988:47).

There are several steps that can prevent defensiveness problems. First, the peer reviewers must be carefully selected with good communication skills. Second, both the peer reviewer and the party being reviewed need to be educated on the peer review process. Last, upper management of the party being reviewed must set a non-threatening atmosphere for the peer review.

Zallen proposes a list of important issues that must be addressed before consensus can be developed on structural design peer review (Figure 4). Before any peer review is initiated, it is important that all members of the peer review team and the party being reviewed completely understand the purpose of the peer review including the scope. The purpose of the peer review may be, for example, to ensure the structural soundness of a new building design. The scope may be limited to only the steel members in the design. The scope defines which structures will be reviewed and what parts of the design are subject to review. From the scope of the review, the qualifications of the reviewers are determined.

- The purpose of the review.
- Which structures will be reviewed.
- What parts of the design will be subject to review.
- Qualifications of the reviewing engineer.

Figure 4. Important Issues (Adapted from Zallen, 1990:209).

Categories. In the last 20 years, engineering organizations developed guidelines for construction-oriented peer reviews. Two categories of guidelines emerged: organizational and project peer reviews (Figure 5). Organizational peer reviews focus on the design organization as a whole. Organizational peer reviews do not review specific projects; they are more concerned with the policies, business practices, and guidelines used by the organization to ensure projects meet standards. Project peer reviews focus on specific project(s) or project management. Project peer reviews focus on design, not the organization in general (American Society of Civil Engineers, 1988).

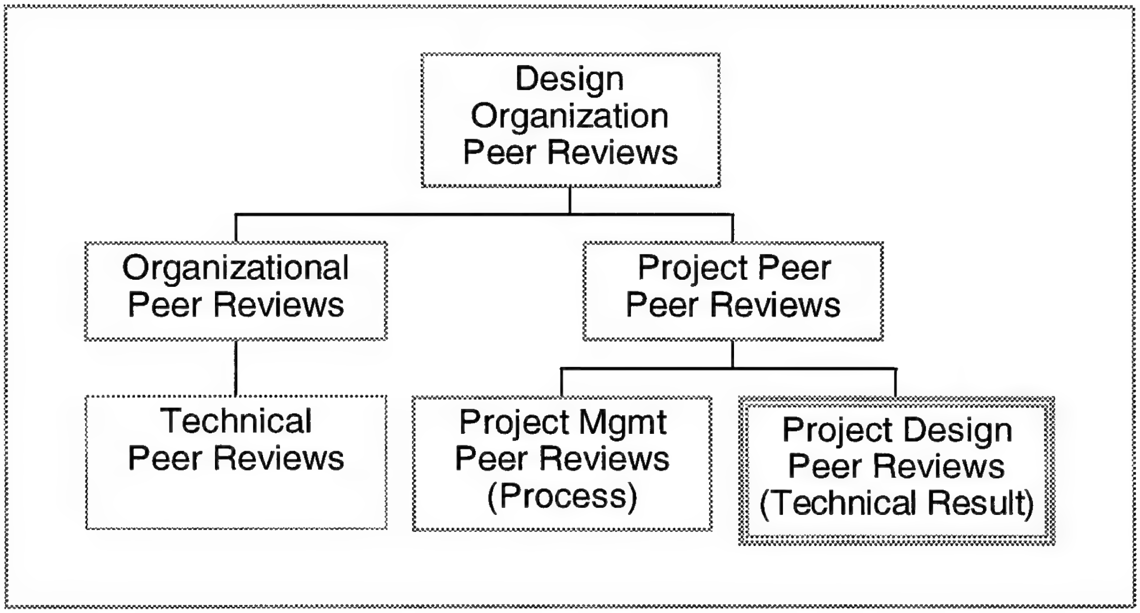

Design Organization Peer Reviews
Organizational Peer Reviews
Project Peer Peer Reviews
Technical Peer Reviews
Project Mgmt Peer Reviews (Process)
Project Design Peer Reviews (Technical Result)

Figure 5. Design Peer Review Categories (Adapted from ASCE, 1988:80; ACEC & CASE, 1988:3)

As an example of an organizational peer review, the Coalition of American Structural Engineers' (CASE) Technical Peer Review was developed in 1986 as a vehicle to limit liability of structural engineers. Technical Peer Reviews could be more appropriately titled "Liability Reduction Peer Reviews" since the primary focus is on ensuring policies, guidelines, and documentation are in place to limit liability. The following are examples of areas reviewed by the Technical Peer Review: contracts and scope of work, project design criteria, calculations and documentation, quality assurance procedures, and office communications (American Consulting Engineers Council, 1988).

There are two general types of project peer reviews: project management peer reviews and project design peer reviews. Project management peer reviews are similar to organizational peer reviews in that they look at the procedures, policies, and documentation supporting design projects. The scope is narrower, however. Project management peer reviews are focused on the layer of management directly above the designers, while organizational peer reviews focus on the management of the design organization as a whole. Project design peer reviews focus on reviewing sample calculations and specific project documentation to verify the adequacy of design concepts (American Society of Civil Engineers, 1988). Since the Air Force's environmental peer review program is project-oriented and it resembles the project design peer review developed by the American Consulting Engineers Council, project peer review is next investigated in detail. Refer back to Figure 5 for the hierarchy of peer review guidance.

Project Peer Review. The American Consulting Engineers Council (1990a) notes distinguishing characteristics of project peer review. First, a peer review is conducted by peers of the project designer or manager under review. A peer is defined as a person of equal or higher technical or managerial standing. For example, if a professional engineer completes the design, then the work should be reviewed by another professional engineer. Second, the peer reviewers should always be independent of the party owning the project under review. Care must be taken to minimize personal interactions between the reviewers and the party under review. Third, a project peer review is an extra effort, not another step in the design process. Fourth, a project peer review can be a one-time event or a chain-of-events with a specific purpose and scope.

The purpose and scope of the project peer review are critical to its effectiveness. The purpose clearly outlines the goals and objectives of the specific peer review and should be clearly communicated to all parties involved in the peer review. The scope focuses the peer review on a specific target giving the peer review direction with a specified duration. The scope should also be clearly communicated to all parties to facilitate preparation and efficiency. A clearly communicated purpose and scope will facilitate the underlying goal of the project peer review--to produce the required results. Further, to ensure the review process is effective, there should be guidelines and metrics incorporated through a quality process (Zallen and Huang, 1992:189).

To clarify what a project peer review is, it is helpful to review what a project peer review is not. A project peer review does not review the management policies, guidelines, nor practices; it is focused on the review of the specific project(s) from a technical perspective. A project peer review is a special event, it goes beyond the standard design review which should still be conducted internally. A project peer review is not intended to validate all design calculations and assumptions, the goal is to verify the design concept and the general pattern of assumptions. To meet this end, it is suggested that reviewers take a "random" sample of the project(s) to review in detail. Because USAF environmental peer review program focuses on the review of the technical aspects of the project, "Project Design Peer Reviews" (American Consulting Engineers Council, 1990a) are most applicable and are described next. Refer back to Figure 5 for the hierarchy of peer review guidance.

Project Design Peer Review. The project design peer review process is generally broken into four sub-processes: planning, document review, on-location review, and last report and follow-up (American Consulting Engineers Council, 1990a). Depending on the requirements defined in the planning process and the adequacy of the document review, the on-location review may not be required. Still, on-location interviews are usually conducted as an added dimension of information gathering.

As with most processes, planning is critical to an effective peer review. The first step in the planning process is to carefully define the requirements and objectives of the peer review. The objectives form the basis for defining the scope of the peer review including the planned timing and duration. The timing of the peer review is often critical to its effectiveness. The peer review must be early enough in the design process to provide input for cost-effective decision making. At-the-same-time, the peer review must be late enough in the design process for adequate data to review (American Consulting Engineers Council, 1990a:7).

The combination of peer review objectives, scope, and duration are used to compose the peer review team. The objectives dictate the specialties required on the team. The complexity and scope of the peer review dictates both the specialties and number of reviewers selected. Precisely defined objectives and scope will increase the efficiency of the planning and the effectiveness of the peer review. Efficiency and effectiveness must be balanced in the peer review process (Chubin and Hackett, 1990:47). Further, increasing the efficiency of peer reviews will help keep down costs; for example, peer reviewers cost the Environmental Protection Agency between $150 and $300 per day (General Accounting Office, 1994a:3).

Document review is a very important process that should be accomplished before the on-location review. Document review includes requesting specific written documents in a specified format. The party being peer reviewed next prepares and sends requested documentation to the peer review team leader. Last, advanced materials are reviewed by the peer review team in preparation for on-location interviews and meetings (American Consulting Engineer Council, 1988:15).

Prior to the on-location meetings and interviews, the peer review team leader should prepare a checklist or agenda of items to be covered consistent with the peer review's objectives and scope. The on-location process can entail a series of interviews, a group interview, or a combination of both. Interviews are used to discover design assumptions and concepts that were not clear in the written documentation. Good communication skills and a non-threatening atmosphere are critical to making the on-location peer review successful (American Consulting Engineers Council, 1990a:8).

The last sub-process in the peer review process is a written and, if desired, an oral report. The American Consulting Engineers Council (1990a; 1990b:4) recommends a written report supplemented with an oral outbriefing. The emphasis throughout the peer review and especially in the final report is on simplicity. The optimum outcome of the peer review is that the assumptions, design concepts, and construction decisions are adequate with no required changes. The American Consulting Engineers Council (1990a:8) goes on further to say that "the team should not attempt to justify its existence by giving a long list of suggestions." Of course, required changes should be documented.

Peer Review Process

The peer review process entails reviewing new ideas for incorporation into the body of knowledge, procedure application, or funding. Following verification, the concepts or designs can be implemented through new scientific procedures, in the case of manuscript review; funding, in the case of grant review; and construction, in the case of construction project review. There are two primary frameworks for performing peer reviews: serial and parallel. Serial systems are usually used for maximum efficiency. The parallel process is used when an extra layer of safety is prudent. There are also three types of blinding that can be used: nonblind, double-blind, or single blind.

All applications of peer review are adopting criteria to guide the process and make it more effective. These criteria establish guidance for peer review process protocol, incorporating risk into the decision making process, and choosing reviewers. The important factors for an effective environmental peer review protocol were discussed in the Project Design Peer Review section. These factors include peer review: objectives, scope, duration, team specialist mix, document review, agenda, and team leader.

Innovation versus Risk Abatement. In furthering the body of knowledge, care must be taken to encourage innovation and to prevent overly cautious behavior, unless dictated by the risks involved. Horrobin describes the interaction between quality control and the encouragement of innovation in peer review of medical manuscripts:

> Peer review can be performed successfully only if those involved have a clear idea as to its fundamental purpose. Most authors of articles on the subject assume that the purpose of peer review is quality control. This is an inadequate answer. The fundamental purpose of peer review in biomedical sciences must be consistent with that of medicine itself, to cure sometimes, to relieve often, and to comfort

> always. Peer review must therefore aim to facilitate the introduction into medicine of improved ways of curing, relieving, and comforting patients. The fulfillment of this aim requires both quality control and the encouragement of innovation. If an appropriate balance between the two is lost, then peer review will fail to fulfill its purpose (Horrobin, 1990:1438).

Horrobin and Gallagher make further observations in regards to the balance between risk and innovation. In academic discussions of this balance, quality control or risk abatement are the dominant goals, while the encouragement of innovation receives little attention (Horrobin, 1990:1439). Gallagher notes that in the 20th century, peer reviewed journals are coming under increased pressure to publish only articles that are 100% verified. This unrealistic expectation undermines the journals ability to further the body of knowledge by publishing cutting-edge research, that is not completely verified (Gallagher, 1989:iv). Regardless of the specific policies of the journal, the goal of a peer review program is to provide the editor with unbiased reviews from diverse specialists in the field. "Thus assisted, the editor can make more broadly based, wiser decisions. The reviewers are more like jurors than detectives" (Gallagher, 1989:iv).

Horrobin states that "innovation is so rare, so valuable, and so central to improvement of patient welfare that innovative articles should be deliberately encouraged and more readily published than conventional ones" (Horrobin, 1990:1439). Horrobin goes on to give eighteen examples of how current peer review encourages conventional articles over innovative ones. One of the examples he gives is on the subject of peer review in the grant-giving process. The review is so restrictive that innovative scientists often tell lies on their grant applications. Similar views have been voiced by at least two Nobel Laureates (Horrobin, 1990:1440).

Peer Reviewer Characteristics. Peer review researchers agree that the specific role of peer reviewers should be carefully determined before the review begins. Cozzens found that while technical experts make excellent peer reviewers, they should not be used to judge non-technical matters. "Scientists are intelligent, and they are expert in their own areas, but this does not make them omniscient" (1987:80).

One pitfall of peer review that Cozzens' case studies of several large organizations uncovered is the importance of establishing required peer reviewer characteristics before beginning the peer review (1987:80). She suggests carefully matching reviewers expertise with the specific project being evaluated. A poor match of reviewer to project will result in a poor quality report. The European Economic Community recommends a panel of six to eight reviewers to ensure adequate coverage of experience and to foster synergy of reviewers (Cozzens, 1987:74).

Care must be taken in selecting peer reviewers. According to "Official Releases: Peer Review Standards Interpretations" (1995:119), "an individual serving as a reviewer must be currently active in public practice at supervisory level in the accounting or auditing function of a firm enrolled in an approved practice-monitoring program." Reviewers must be current in the field reviewed. Atkinson found that people that are well-known for innovation do not necessarily possess current competency in leading-edge technology (1994:151).

Taubes (1993) and Stossel (1985) find that high positions are a handicap and that reviewers under 40 years old are the best. High administrative and academic rank is a handicap (Taubes, 1993:26). This is likely attributable to the higher rank positions requiring more time, allowing less time to be devoted to reviewing manuscripts. Taubes (1993:26) hypothesizes that top ranking reviewers that are highly published did not spend adequate time to thoroughly review, while more junior reviewers tended to spend more time on the review, and it showed. Further, the best reviewers tend to be under 40 which could also be due to higher responsibilities with increased age resulting in less time for reviewing (Taubes, 1993:26).

Beyer et al (1995:1243) finds that authors submitting manuscripts for peer review fared best if they were male, associate professors, and had funding. Male bias is common in many fields and is not a surprising finding in the scientific community, although as with other areas, steps must be taken to minimize this bias. The associate professor bias is less intuitive. Perhaps associate professors are more successful because their careers are more dependent on the success of the research. Having funding indicates a researcher has been successful in the past competing for funding and is considered more qualified by having passed this litmus test.

McCabe (1991:111) cites the important qualifications for a good CPA reviewer (Figure 6). A good reviewer meticulously investigates the firm's quality control system which helps fine-tune and lend credibility to the system under review. Reviewers who have in-depth experience bring to the review a wealth of knowledge that can be applied to help improve the reviewed process. McCabe (1991:111) recommends a minimum of five years experience in accounting and auditing. Reviewers should have good listening skills and should be selected for their ability to actively listen (McCabe, 1991:111). Good people skills help the reviewer uncover details not found in written reports much more quickly and accurately. Lastly, a good peer reviewer is dedicated to the reputation of the profession.

1. Meticulous Attention to Detail
2. Experience
3. Good Communication Skills
4. Excellent Interpersonal Skills
5. Dedicated to the Profession

Figure 6. Traits of a Good Reviewer
(Adapted from McCabe, 1991:111)

Peer Reviewer Evaluation Instrument. Although there have been several studies on the quality of manuscripts, the effects of blinded studies, and the reliability of peer review, there have been few studies on the effectiveness of grading instruments used by editors to evaluate the manuscript reviewers. Feurer et al. (1994:98) investigated the effectiveness of a grading instrument used to evaluate the quality of peer reviewer's analysis of manuscripts. A grading instrument could be used by editors to determine which reviewers are "consistently helpful, outstanding or weak, timely or late, and exacting or lenient, and those who contribute the most and least to the peer review process" (Feurer et al, 1994:98).

If reliability is used as a metric of the effectiveness of peer review, Feurer and colleagues' evaluation of a grading instrument for peer reviewer reliability is a good measure. The grading instrument uses seven attributes arbitrarily determined to be important for the quality of peer review: timeliness, grade sheet, etiquette, section-by-section review, other supporting references, summary and/or recommendations, and new insights/perspectives. Forty-one reviews of 23 manuscripts were analyzed using these seven attributes. For example, if the grading sheet was filled out by the reviewer, the review received one point, if not, no points were awarded. Likewise, for the section-by-section review, if the review addressed each section, the review received 3 points on a scale of 0 to 3.

The study found "the interclass correlation coefficient was .84 ($p < .001$) and a lack of difference between mean scores was demonstrated by analysis of variance ($p = .46$)" (Fuerer et al, 1994:98). This indicates that reliability is high for the peer review process, assuming the validity of the instrument for measuring reliability (Shrout and Fleiss, 1979).

The determining factor for the instrument's ability to accurately measure reliability is if the instrument measures attributes that are universally held by journal editors to be important. Since this experiment was done to develop an instrument for a specific journal, developing the criteria using the editorial staff's perspective is acceptable. For this instrument to be applied universally to other journals or other media, the attributes of the instrument need further evaluation, perhaps by a survey to rank the importance of attributes.

Summary

Tracing the historical development of peer review indicates a trend toward non-systemic maturation of programs. Peer review, especially in early journals, was compartmentalized to a large degree. Grant and accountancy peer review programs developed more systematically due to cross-feed of information. The recent development of project design peer review included a high degree of sharing between different professional organizations.

The reliance on emerging technology in the environmental field is similar to issues found in scientific manuscript, grant, accountancy, and research program review. Important issues of the peer review processes include the balance between caution and innovation, the parameters of the process, and the selection of reviewers. Through a review of these similar peer review processes, lessons can be gleaned to enhance the current Air Force Peer Review Program.

The majority of the literature that pertains directly to environmental project design peer review comes from professional engineering organizations. The guidance documents written for project design peer reviews are directly applicable to the environmental field. The primary difference lies in the fact that the environmental field is still an emerging science, in contrast to the core engineering fields such as structural engineering. While structural engineering has books dedicated to specific calculations for truss members, for example, there are no specific engineering books for the environmental field. Through a review of the varied application of peer review in related fields, this research attempts to apply the lessons learned in other fields to the Air Force's peer review process.

Chapter 3

Survey Development

Construct Development

Since the underlying objective of this research is to quantify peer review effectiveness, the effectiveness construct is first grounded in Air Force policy to establish the requirement of the criterion construct. Next, the constructs that define the peer review process effectiveness are discovered. These constructs are grounded in the academic literature, industry standards, peer review process observations, and issues from interviews. Each construct is described using triangulation to combine the data sources. Following their description, the predictor constructs are grounded in Air Force policy.

Effectiveness Construct. Many of the authors in the peer review literature look at the parameters of the peer review process (Hirschman, 1994; Knoll, 1990; Luce, 1993; Atkinson, 1994; Gallagher, 1989; Cicchetti, 1991, Daniel, 1993; Oxman, 1991). The importance of effectiveness versus efficiency is frequently debated.

The environmental peer review process balances efficiency with effectiveness. Efficiency is the rate at which the project peer reviews are completed, while effectiveness is the degree to which the peer review meets the program objectives. The current Air Force environmental peer review process focuses on effectiveness due to the the complexities and costs of the projects being a tighter constraint than the number of projects.

Two measures of peer review effectiveness, proposed by Luce (1993), are result oriented. The first measurement is the degree to which the review decreases the number of modifications of decisions that are later deemed correct. The second measurement is the increase in number of modifications of decisions that are later deemed incorrect. Unfortunately, these output measures may take many years to appear in environmental projects. This research attempts to define constructs that are process metrics of peer review effectiveness.

Environmental project peer review is used to ensure the Air Force is meeting its requirements as established by legislation, regulations, and Air Force policy. The Air Force has established four objectives (Figure 7) to meet the requirements established by the federal acts and executive orders discussed in the Introduction Chapter.

1. Identify sites and areas of concern;
2. Investigate threats;
3. Cleaning up sites; and
4. Closing out IRP sites.

Figure 7. Air Force Environmental Program Objectives (DoAF, 1993b:1-1; DoAF, 1994b:1)

Figure 8 shows the nine CERCLA criteria for meeting the cleanup objectives in Figure 7.

In 1995, the Air Force Air Staff delegated the responsibility of developing and

1. Cost;
2. Protection of human health and the environment;
3. Short-term effectiveness;
4. Compliance with applicable or relevant and appropriate requirements;
5. Long-term effectiveness and permanence;
6. Reduction of toxicity, mobility, or volume;
7. Implementability;
8. State acceptance; and
9. Community acceptance.

Figure 8. Air Force Environmental Program Criteria (Adapted from DoAF, 1993b:1-1)

maintaining the Air Force environmental restoration project peer review program to the Air Force Center For Environmental Excellence (AFCEE). See Figure 9 for the Air Force Center for Environmental Excellence four-fold purpose for the Air Force Peer Review Program (AFCEE, 1996:1).

1. Validate the technical merits of the proposed remedial actions to achieve the stated remediation goals.
2. Evaluate the adequacy of the rationale used to set remediation goals.
3. Ensure incorporation of a properly-conducted, risk-based approach as a remediation decision tool.
4. Provide validation documentation to accompany funding requests and/or technical recommendations to improve proposed remedial actions.

Figure 9. Four-Fold Peer Review Program Purpose

Table 1 shows how this four-fold purpose meets the nine criteria. The right hand column of Table 1 shows which of the four purpose statements ensures the specific criteria is met. These nine criteria (Figure 8) ensure that the program satisfies the four objectives of Air Force environmental restoration policy (Figure 7). The peer review process is deemed effective if it meets the four-fold purpose (Figure 9). For years CERCLA projects were based on the first seven criteria in Figure 9. Only recently were state acceptance and community acceptance included. State acceptance was included after the Air Force deemed it was appropriate to include regulatory input; the Air Force retained lead responsibilities and did not require regulatory input. Community acceptance, however, was included to increase participation by members of affected or potentially affected communities in the cleanup process at military installations to meet the requirements of CERCLA (Trevino, 1996).

TABLE 1. HOW PEER REVIEW PURPOSE MEETS NINE CRITERIA

Nine Criteria	**How Four-Fold Peer Review Purpose Meet Criteria (Purpose #)**
1. Cost;	by validating funding requests (4)
2. Protection of human health and the environment;	by ensuring risk-based approach (3)
3. Short-term effectiveness;	by validating technical merits (1)
4. Compliance with applicable or relevant and appropriate requirements;	by evaluating adequacy of rationale (2)
5. Long-term effectiveness and performance;	by validating technical merits (1)
6. Reduction of toxicity, mobility, or volume;	by validating technical merits (1)
7. Implementability;	by validating technical merits (1)
8. State acceptance; and	by validating adequacy of rationale (2)
9. Community acceptance	by evaluating adequacy of rationale (2)

Construct Discovery Method. The literature, observations, interviews, and feedback forms are next reviewed for characteristics which describe the seven environmental project peer review constructs. First, the methodology for defining the seven environmental project peer review constructs is described. Then, the following seven constructs are defined: focus, agenda, facilitator, written preparation, oral presentation, peer review team characteristics, and peer reviewer characteristics.

The naturalistic inquiry paradigm is used for data collection. Patton (1990:40) defines naturalistic inquiry as the study of "real-world situations as they unfold naturally; [it is:] non-manipulative, unobtrusive, and non-controlling [and] open to whatever emerges. [There is a] lack of predetermined constraints on outcomes." Naturalistic inquiry uses data from three basic sources of qualitative data collection: (1) interviews, (2) observation, and (3) written documents (Patton, 1990:10,118).

Data triangulation--the use of a variety of data sources in a study--is recommended to increase the reliability of findings (Patton, 1990:187). In this research, literature review, observation, interviews, and feedback forms are used to collect data. Literature review is the primary source of background data used to develop the constructs. The findings from the literature are validated using the other data sources. Using a combination of data sources and collection tools allows for a cross-check and validation of findings (Patton, 1990:244). The methodology for each of these data sources is described later in this chapter. These data are analyzed using similar methodology.

The primary method used to reduce and interpret qualitative data in this research is content analysis. There are many methods used for document content analysis (Creswell, 1990:154-155). The method used in this research was based on the steps proposed by Tesch (1990:142-145). The resulting method is a six step approach in Figure 10.

1. Get a sense of the "big picture."
2. Pick one document (interview) and read for main topics of information.
3. Make a list of main topics from several documents.
4. Using this list, go back over all the documents.
5. Find the best wording for categories and attempt to merge the categories.
6. Assemble data belonging to each category and perform a preliminary analysis.

Figure 10. Iterative Content Analysis Steps. (Adapted from Tesch, 1990:142-145).

While the data sources are described separately in this section, the collection, reduction, and analysis was necessarily combined due to the overlap of the data collection. For example, interviews were conducted during peer review observation due to the limited availability of peer review experts. At the same time, these peer review experts also completed feedback forms following the peer reviews.

Literature Review Content Analysis. The literature review followed the content analysis methodology described in Figure 10. First the literature was reviewed for the "big picture." Because the literature is deficient in project peer review process methodology, the picture was built from reviewing related fields such as manuscript refereeing, grant peer review, and accountancy peer review. Further, the majority of related academic research is in the field of manuscript refereeing; the tools developed by and for journal editors for peer review was used as a basis for lessons learned.

The lessons learned were used for generating ideas for process improvement and for developing categories which became the constructs to predict peer review effectiveness. The deficiency of related project peer review process literature required a literature review that generated ideas from related fields.

Observations. Similar to the literature review, the first step in observation data collection was to paint the "big picture" (Figure 10). The data collection was performed from 28 February 1996 through 2 March 1996 on a trip to observe the environmental project peer review process in San Antonio, Texas. Peer reviews at both a closed base in California and a base on the closure list in Texas were observed. (Appendix B). Observation methodology was applied to the data collection at these peer reviews.

There are four types of observation that are recommended by Creswell (1994:150): "complete participant--researcher conceals role; observer as participant--role of researcher is known; participant as observer--observation role [is] secondary to participant role--and; complete observer--researcher observes without participating."

Patton (1990:217) describes five areas of variation in observation methodology. The first is the role of the observer which varies from a participant with a concealed researcher role to a complete observer. The complete observer methodology was used in this research. The second area of variation is the covertness of the researcher. The range is from totally overt to totally covert. Since complete observer methodology is used, totally overt methodology is used to facilitate clear observation.

The third variation in methodology is the degree to which the researcher's purpose is conveyed to those under observation. Because the research is non-threatening, more could be learned from an open explanation of the purpose. By explaining the purpose of the research, those under observation are more likely to be less suspicious of the researcher resulting in a more accurate case study. The fourth variation is the duration of the observations. Due to limited funding and timing of the peer reviews, observation was limited to one-full day. The last variation in observation methodology is the focus of observations. Because the intent of this research was focused on the entire peer review process, the focus of the observations was, likewise, broad in scope.

The observation methodology used was modeled after suggestions made by Patton (1990:219-244). The data collection was focused at a low level of detail to capture the entire peer review process. Observation included details such as: the setting, verbal and non-verbal statements, planned and non-planned activities, program documents, and what did not happen. Field notes were detailed and concrete; Patton (1990:240) notes that field notes are often vague and overgeneralized.

The observation notes were reviewed to complete steps one through five of the content analysis procedure (Figure 10). Special attention was given to areas that were not addressed in the literature such as facilitator and reviewer characteristics. Due to the restraints of time and the availability of experts, interviews were often conducted simultaneously with peer review observation.

Interviews. Creswell (1994:150) describes three types of interview data collection methodology: "Face-to-face--[researcher is] one-on-one [with subject], in person interview; telephone--researcher interviews by phone; and group--researcher interviews informants by group." Because the face-to-face interview methodology provides the highest detail of information, it was used when possible. When face-to-face interviews were precluded by funding or timing, telephone interviews were used. For example, telephone interviews were used at the beginning of this research before sponsorship funding was available. See Appendix C for notes on interviews used in this research.

Interviews were used first in this research to get the "big picture." Air Staff and major command peer review program managers were contacted by telephone and interviewed for ideas. Face-to-face interviews were conducted when possible for peer review experts on Wright-Patterson Air Force Base. Next, face-to-face interviews were conducted at the Air Force Installation Restoration Management (AFIRM) Review from 31 January 1996 through 2 February 1996. These interviews began to go beyond the big picture and to fill data gaps to develop the main topics of information (Figure 10).

Face-to-face interviews were continued on the Base Realignment and Closure project peer review data collection trip. Peer review experts were interviewed during observation and after the peer reviews were completed. The purpose of these interviews was to fill gaps and to complete the content analysis process. Further, interviews were conducted to begin development of the quantitative survey discussed later in this chapter.

Open-Ended Feedback Forms. The open-ended question feedback form was developed to capture lessons learned and ideas to improve the peer review process (Appendix D). This form gives general instructions, "Please use this page to make comments and lessons learned ideas which will assist us next year to improve the Peer Review." This statement is general enough to minimize predisposed response bias (Patton, 1990:295). These forms were given to peer reviewers in the beginning of the peer review process and were collected upon process completion. This allowed reviewers the latitude to take notes on the form as ideas arose.

The free-form responses on these forms was codified using the content analysis methodology (Babbie, 1990:45). This methodology was not applied to the responses in isolation to the other data sources. The constructs, or categories, that were derived from the literature review, observations, and interviews were used to codify and further refined with the free-form responses on the open-ended question feedback forms.

There are several pros and cons of open-ended questions (DoAF, 1993:12-13). One positive attribute is that open-ended questions give respondents latitude to answer in their own words. Further, since the data being requested were new ideas and lessons learned, the exhaustive categories were difficult to develop. Collectively exhaustive categories are required for well structured closed-ended questions. The negative attributes of open-ended questions is they must be codified, or categorized, for content analysis. This process involves taking infinitely variable responses and fitting them into discrete categories. If possible, it is often more efficient to spend the time up front in planning and establishing categories for closed-end questions. Unfortunately, the creative responses requested do not fit easily into a closed-ended question format.

Data Triangulation. Data triangulation was used to cross-check and validate the results to synthesize the refined constructs of the peer review process, see Figure 11. Based on the results from the data sources, the wording of the constructs was refined and the number of constructs was minimized by category merging. The primary source of

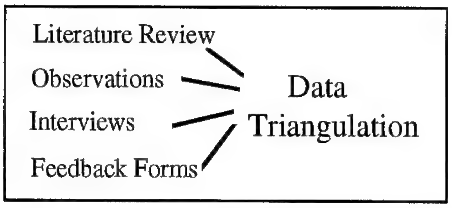


Figure 11. Data Triangulation

data for developing the constructs was literature review. The other data sources were used to ensure the findings from the literature were appropriate for the Air Force application.

Seven Predictor Constructs. This section introduces and describes seven constructs that define peer review effectiveness. There was little academic literature on peer review effectiveness; the majority of this section is based on professional society guidelines on project design peer review. These guidelines were developed substantially from field experience without the aid of academic study. As a result, the characteristics defining the seven constructs are grounded largely on field experience with little reliance on hypothesis testing experiments. Observations, interviews, and feedback forms are used to validate the literature review findings.

Focus. The focus construct is composed from two general categories: objectives and scope. The scope of the environmental project peer review should be clearly defined. Through observation of a peer review of projects from a closed Air Force Base in California (Appendix B), it was evident that the object and scope of the review are sometimes weakly defined and/or communicated. Zallen (1990:212) argues that the peer review is "not intended as a comprehensive design check; it is only intended to verify that the design is conceptually correct and that there are no major errors." Gallager states that "The reviewers are more like jurors than investigators." The scope of the review should be focused on verifying the facts as presented.

While the results should be taken at face value, it is still appropriate to check that the results lie within the expected range. Of the two Air Force Base Closure Agency peer reviews observed (Appendix B), the peer review teams measured project values against a range of reasonable values for cost, risk, and other project values. Sample calculations were also performed. Zallen (1990:213) states that sample calculations are appropriate if within the scope of the peer review. As with the legal system, the underlying theme of the review must aim at having the burden of proof lie with the reviewers.

The Project Peer Review Guidelines (ACEC, 1990:3) places heavy emphasis on the importance of clearly communicating the peer review objectives and the scope. The objectives of the peer review should go beyond the normal checks and balances applied throughout the design phase. The objectives should focus the peer review on validation of the assumptions and design concepts. Stating the desired end-result of the peer review with precision allows the review team to focus on the subject quickly and efficiently and minimize review time (ACEC, 1990:7).

ACEC (1990) recommends that the peer review focus on validating the design criteria, assumptions, and concepts. The design criteria should be validated against all applicable local regulations and the owner's [Air Force's] objectives. The design assumptions should be validated to ensure they encompass the design criteria and all major aspects of the project. Last, the design concepts should be validated to ensure they are technically sound and meet or surpass the design criteria and assumptions (ACEC, 1990:16).

While the design criteria, assumptions, and concepts are reviewed, special attention should be given to non-standard designs (ACEC, 1990: 24). All non-standard assumptions and designs should be reviewed to ensure they are technically feasible and cost-effective. If applicable, sample calculations should be worked separately to check non-standard designs (ACEC, 1990:18; CASE, 1988:39).

The importance of a focused peer review process also surfaced through observations, interviews and feedback forms. While the primary source for focus construct data is the literature, both observations and interviews were used as a check to ensure the literature findings were applicable to the Air Force application.

Appendix B discusses a peer review of a closed base in California in which a clear focus was lacking. A clearly communicated focus through the aid of an agenda could have made the process more effective. The Air Force Base Conversion Agency and the Air Force Center for Environmental Excellence attempted to explain the focus of the peer review in pre-peer review briefings each morning. From comments made during the peer reviews, the focus and goals of the peer review were not clearly communicated from these meetings. During the peer review, several people made comments that they did not know what the purpose of the peer review was. The facilitator repeatedly attempted to keep the peer review focused on choosing the best technology.

While the environmental project peer review does include a review for cost savings, the literature recommends focusing on validating the design (ACEC, 1990:4; ASCE, 1988:79). One of the primary goals of the AF Environmental Peer Review Program is to ensure proposed actions are cost appropriate (Air Force Center for Environmental Excellence, 1996:5). While it is important to ensure funding estimates are validated for the established requirement, it is equally important that the peer review does not primarily focus on cutting costs. The primary focus should be on determining, "Is it [the design] good enough? (ASCE, 1988:79).

A facilitator can assist in keeping the peer review focused on the objectives. The importance of keeping the process on track and focused was mentioned in two interviews (Appendix C). In both interviews, a facilitator was recommended to ensure the peer review is kept focused. Further on the feedback forms in Appendix D, both regulatory and funding issues were discussed in terms of focus.

Effective design project peer reviews have a well-defined focus. Since design projects cannot be economically reviewed in detail, a scope must narrow the review to a representative sample of the project design (ASCE, 1988:81). Quality in the Constructed Project (ASCE, 1988:83) recommends that the scope be defined in detail before the peer review begins; the scope should specify: what parts of the project will be reviewed, the review process that will be followed, and the agenda of the peer review. The scope of the peer review should be stated with precision (ACEC, 1990:7).

Agenda and Facilitator Constructs. An agenda and a facilitator are very useful tools for keeping the peer review focused on the objectives. The two Air Force Base

Closure Agency peer reviews observed had a facilitator, but neither had a formal agenda. The facilitator was an asset and lacking a formal agenda appeared to be a handicap.

A precursor or compliment to an agenda is a checklist of items to be reviewed. A checklist of questions and additional information for on-site peer reviews is recommended by many of the professional organizations with review programs (American Consulting Engineers Council and American Society of Civil Engineers, 1990a; Coalition of American Structural Engineers, 1988; Association of Soil and Foundation Engineers, 1990). A formal agenda or schedule is also recommended by these professional organizations.

Both agendas and facilitators were deemed important in the research observations, interviews, and feedback forms. A clear agenda was not used in the observed peer review processes (Appendix B). The lack of a pre-published and clearly defined agenda was a hindrance to the effectiveness of the observed peer reviews; however, a strong facilitator at the peer review for a closed Air Force Base in California compensated for the lack of structure through dynamic "people skills."

The Project Peer Review Guidelines (ACEC, 1990:8) recommends tailoring the checklist (or agenda) to the specific objectives and scope of each individual peer review. While general checklists are an aid to efficiency and consistency, narrowing the focus of checklists will save time and will make the peer review more productive. Further, checklists can aid the reviewers in covering all non-standard areas of the projects (ACEC, 1990:24).

A peer review team should have a facilitator, or at a minimum a team leader, to keep the peer review focused on the objectives and scope. Many of the professional organizations reviewed with a team of peer reviewers recommend establishing a team leader (ACEC & CASE, 1988:33; Department of Energy, 1991:4; ASFE, 1990:6). The team captain is responsible for establishing and maintaining open communication between the reviewee and the peer review team. Through the use of checklists and an agenda, the team leader is also responsible for ensuring all items that are intended to be reviewed are adequately considered (ASFE, 1990:31).

Two interviews dealt with the importance of an agenda and one stressed the value of a strong facilitator (Appendix C). A clear agenda facilitates preparation of both the peer review team and the project managers. When an agenda is not feasible, a strong facilitator can focus the peer review and help prevent an ineffective peer review.

Two of the feedback forms (Appendix D) implied the value of an agenda and three discussed the need for a facilitator. The peer review should be structured with time allocated for overview of projects and breaks. The facilitator is a critical part of the peer review and, in several cases, was instrumental in performing an effective peer review.

An agenda helps keep the on-site peer review focused on the objectives and the scope. A focused peer review results in an efficient use of time. To further increase the efficiency, the agenda should be clear and easy to follow. The team leader and reviewee mutually develop the agenda based on the peer review's objectives and scope (CASE, 1988:41). The reviewee should be allotted time within the agenda to place all the projects in context. The reviewee should be allotted additional time to introduce each individual project (Department of Energy, 1991:11).

Written Preparation. Written document review should be an integral part of project design peer review (ACEC & ASCE, 1990:17; ASCE, 1988:85; ACEC & CASE, 1988:15; DoE, 1991:9). Written documentation was used in all observed peer reviews. Many of those interviewed (Appendix C) stated they thought the current written preparation was not effective. Peer review team members should be sent project documents prior to the review. These documents should be of adequate detail to generate questions for the on-site portion of the peer review. Additionally, reviewers should receive project documentation with sufficient time for complete review.

The importance of written preparation was strongly emphasized through observation, interviews, and feedback forms. Approximately six inches of written materials were given to the peer review team for the peer review from a closed Air Force base in California with inadequate time for a thorough review (Appendix B). Although there was written guidance outlining the written requirements (AFCEE, 1996), the closed base did not utilize the guidance nor did they accept in-person assistance which was offered by the Air Force Center for Environmental Excellence. The documentation was so immense that it was practically of no use. From observation, good organization and concise writing is of high importance for useful written preparation.

Two of those interviewed expressed concern for an improvement in both the quality of written material and the degree to which the material is read by reviewers (Appendix C).

Several factors hindered the reading of prepared documentation. First, the written documentation was too thick to be easily read. Second, the written material was not distributed to peer reviewers with adequate time to prepare.

The importance of written preparation is mentioned in 9 out of 23 feedback forms (Appendix D). Written documentation is important to communicate the maturity of the project. Written communication is essential for developing an agenda and advance questions for project managers. Written preparation was a common weakness in the peer review process.

To facilitate peer review efficiency, ASCE recommends reviewers thoroughly check project documentation prior to the on-site review (ASCE, 1988:85). The review of documentation begins when the peer review team leader requests specific project documentation from the reviewee. These documents are intended to focus the reviews on the stated objectives and scope of the peer review. The reviewee next submits a package of documents to the team leader who distributes them to the peer review team (DoE, 1991:3). Last, "Each reviewer is responsible for reviewing the advance material and becoming familiar with it" (ACEC & CASE, 1988:15).

Oral Presentation. Oral presentation is the most strongly voiced peer review process characteristic throughout the observations, interviews, and feedback forms. During the peer review of a closed base in California, the project manager's sale of the projects was addressed as important (Appendix B). The project manager did not clearly communicate the overview of the projects nor the specifics for review. Further, the project manager conveyed an adversarial attitude that was evident through the oral presentation. These two factors negatively impacted the effectiveness of the peer review. A clear and concise presentation of projects coupled with a positive, non-critical attitude were lacking in the peer review and were a hindrance.

Improvements for oral presentation were explained in detail through two interviews (Appendix C). An overview is first recommended to ensure all peer reviewers have a common knowledge base of the general aspects of the base remediation program. Next, the specific projects should be introduced; the critical aspects of the projects should be highlighted.

The Department of Energy (1991:11) recommends the reviewee take approximately 15 minutes to place all related projects in context. This portion of the briefing is at a low level of detail to "paint a picture" of all related projects under review. As each individual project comes up for review, the reviewee, takes approximately five minutes to introduce the specific project. This briefing is a lower level of detail and focuses on the areas specified by the objectives and scope of the peer review.

The oral presentation should be at an appropriate level of detail that facilitates the effectiveness of the peer review but is not too detailed to loose efficiency. Twelve out of twenty-three feedback forms expressed concern for a clearly delivered oral presentation. Oral presentation was the most frequently mentioned area on the feedback forms. Oral presentations were highlighted as a critical area and an area that would benefit from clear guidance that is standardized for clarity.

The oral presentation and interviews tell the peer review team details that are not readily apparent from written documentation (ACEC, 1990:8). During these interviews, the reviewee provides additional written documentation to supplement both the interview and the advance written documentation package. These discussions may either confirm or reverse the preliminary results from the advance written documentation (ACEC & CASE, 1988:17). For an effective and efficient interview, it is important that the reviewee has information readily available that is requested by the review team.

Peer Review Team. Even as early as 1901, peer review program administrators saw the importance of having the correct mix of peer reviewers. In Moyer's account (Burnham, 1990:1326), he learns that having a chemist on the peer review team would have saved a lot of unnecessary effort. Like journal peer review, specialization has made it difficult at best for one environmental manager to have the knowledge required to review an entire remediation project. In environmental project peer reviews, the correct mix of specialists is critical.

Having the proper mix of specialists on the peer review team surfaced during the research observations, interviews, and in reviewing the feedback forms. Two issues were emphasized: duplicity of specialists and the lack of a cost analyst. One technical expert from each specialty such as chemistry is sufficient. One of the people interviewed stressed that three chemists was probably overkill. Since the current focus of the Air Force Base Conversion Agency peer review process is to validate funding, a cost analyst is mentioned as a necessary addition to the peer review team.

Ciccetti (1993), Luce (1993), and Sinclair (1993) debate the appropriate balance between rater consistency and variance. Luce and Sinclair both argue that increasing variance should be the goal of peer review because it will maximize the range of errors caught and improvements made in the peer review process. Ciccetti argues that reviewer consensus is needed to validate the findings of the review. Since environmental remediation often involves engineering decisions without generally accepted solutions, striving for consensus is not usually practical. Where the scope of the remediation is limited and the engineering decisions are generally accepted, seeking consensus is beneficial. On the other hand, when the scope of the remediation project is broad and engineering decisions are not generally accepted, the peer review process would benefit from a diversity of reviewers with divergent professional perspectives.

The size of the peer review team and the mix of specialists on the team depends on the objectives and scope of the peer review and the complexity of the project design under review (ASCE, 1988:83). For large, complex design projects most professional organizations suggest at least two and ideally between five and nine peer reviewers on the team (ACEC & ASCE, 1990:9; DoE, 1991:3; Cozzens, 1987:74). The team should be composed to facilitate an objective and constructive evaluation of the project under review. This is achieved, not only through the careful selection of individual peer reviewers, but also through the size and mix of specialists on the peer review team.

The mix of specialists on the peer review team is important for an effective and efficient review. For large, complex reviews, the DoE (1991:3) includes reviewers from industry, academia, research laboratories, the government, and other appropriate sources.

Each team member brings unique experiences, educational background, and insights to the peer review team (ACEC & ASCE, 1990:8). These qualifications must be balanced to cover all objectives within the entire scope of the review. Project Peer Review Guidelines (ACEC & ASCE, 1990:10) recommends the team be composed of registered professionals with a minimum of 15 years experience to lend credibility and ensure adequate qualifications for reviewing. While this level of experience is not possible for all reviews, it is important that professional experience is heavily considered when composing the peer review team.

Peer Reviewer. The individual peer reviewers must be selected based on a wide-range of qualities and qualifications. The reviewers must have good communication skills to facilitate information gathering interviews and be very knowledgeable in the reviewed field of study. Additionally, the reviewers should be adequately independent from the reviewed projects to ensure objectivity. By listing the important skills for a specific peer review, the reviewers can be carefully selected to advance the effectiveness of the peer review.

Listening, interviewing, diplomacy, and the exercise of good judgment are important communication skills that peer reviewers must possess (ASCE, 1988:83; ACEC & CASE, 1988:64; McCabe, 1991:111). Reviewers must be active listeners that pay close attention to detail; active listening is listening with a heightened sense of awareness and focusing on the person speaking. In addition to active listening, reviewers must have sharp interview skills that enable them to get to the "root of the problem." Further, the reviewers should be sufficiently independent of the projects under review to prevent loss of objectivity and to aid in diplomacy and the exercise of good judgment (ACEC & ASCE, 1990:3). Throughout the peer review process, diplomacy and good judgment lend credibility to the process and improve the effectiveness of the review.

The importance of peer review team members with good "people skills" was apparent through observation and through content analysis of the feedback forms. Peer reviewer and facilitator people skills were instrumental in making the peer review process at a closed base in California effective despite an adversarial climate (Appendix C). The importance of professionalism is also mentioned on one feedback form (Appendix D).

The qualifications of the individual reviewers must lend credibility to the process. The reviewers must possess equal or greater qualifications than those designing the projects under review (ASCE, 1988:78). The team members should be senior design engineers or scientists that are well respected by management and the design team of the projects under review (ASCE, 1988:83). Peer reviewers must be carefully selected to ensure they have adequate communication qualities and qualifications to lend credibility to the peer review process and promote review effectiveness.

Seven Predictor Constructs Grounded in Air Force Policy. The previous section described how sources external to Air Force policy define the characteristics which compose the seven peer review process predictor constructs. This subsection describes how the seven constructs--focus, agenda, facilitator, written preparation, oral presentation, peer review team characteristics, and peer reviewer characteristics--are defined by Air Force policy through the Peer Review Guidance Document (AFCEE, 1996).

The focus construct measures how well the peer review is kept targeted at the four-fold purpose of the Peer Review Program. The construct measures how well the peer review focuses on the (1). the adequacy of rationale used; (2) whether a risk-based approach was used; (3) validating the technical feasibility; and (4) validating the funding estimates to ensure the project is cost-appropriate (AFCEE, 1996:1).

The agenda and facilitator constructs measures how closely the peer review follows the planned agenda and how well the facilitator kept the peer review on time and focused on the purpose (AFCEE, 1996:3). The guidance document (AFCEE, 1996:6-7) gives a sample agenda emphasizing its importance to the peer review process. The facilitator's role is described and required for each peer review (AFCEE, 1996:3).

The Peer Review Guidance Document (AFCEE, 1996) specifies the written documentation that is required in advance of the peer review. The Peer Review Information Package contains information describing: (1) the site characterization, (2) the risk-based approach, (3) the proposed technology application, and (4) non-technical concerns such as Federal Facilities Agreements and Native American concerns (AFCEE, 1996:4-5). The Peer Review Information Package is required from the reviewee at least four weeks prior to the peer review to allow sufficient time for distribution and complete examination by the peer reviewers. The Peer Review Information Package is forwarded by the Peer Review Team Chief to the individual reviewers at least two weeks prior to the peer review (AFCEE, 1996:6).

The Peer Review Guidance Document (AFCEE, 1996:6) requires an oral presentation of projects being reviewed. The reviewee is required to give an overview of each project under review; multiple sites are briefed separately. The questions that should be addressed by the reviewee are outlined in detail (AFCEE, 1996:9).

The peer review team is clearly specified (AFCEE, 1996). There is a team leader, facilitator, recorder and specialists as required by the project under review. Typical specialists include: chemists, hydrologists, geologists, risk assessors, and lawyers. The peer review team members are requested from organizations separate from the organization representing projects under review to maintain objectivity.

Specific peer review characteristics were not addressed in the Peer Review Guidance Document (AFCEE, 1996). While it appears intuitive, on the surface, how to select reviewers, guidance in this area would be constructive. It is important that selections are not based solely on the professional qualifications. Qualities such as good communication skills and being respected by management and peers are also important selection criteria.

Final Phase of Survey Development

Survey Cover Letter and Instructions. The cover letter and instructions were developed to maximize the response rate to the survey (Appendix E). A Privacy Act statement was not required because neither the respondent's social security number nor their name was solicited. The surveys were tracked by putting a number on each survey and recording the number on the distribution list.

The cover letter was written to be concise, informative, and motivational to encourage response. The cover letter was signed by the section chief, who was directly responsible for implementing the Air Force peer review program guidelines (AFCEE, 1996:1). The purpose of the survey was described quickly and the confidentiality of respondents was assured (DoAF, 1993:29-30).

The cover letter was followed by an instruction sheet. The instruction sheet directs the person completing the survey to base their responses on a particular peer review on a particular day. The instructions also give a deadline for responses and a fax and phone number for questions or correspondence. The survey items follow the instruction page. Appendix F discusses the process used for survey item development. Appendix G discusses the procedures followed for survey approval.

Pretest Instrument. Upon completion of item refinement and survey design, the survey was pretested. The purpose of the pretest was to get feedback on how well the cover letter motivated, how clear the instructions were, and the clarity of the items. The Air University Sampling and Surveying Handbook (DoAF, 1993:34) recommends choosing a small group of three to ten to complete the pretest.

Seven people were selected to pretest the survey based on their qualifications. One layman, outside the Air Force, was selected to ensure the language was understandable and free of jargon. Two Air Force Institute of Technology (AFIT) survey experts reviewed the questionnaire to check for survey item wording and survey format. Three people who were involved in the environmental project peer review program reviewed the survey to ensure all items were relevant. Last, one AFIT professor reviewed the survey for professional clarity.

Chapter 4

Method

Research Design

The objective of the survey was to measure perceptions of the peer review process. The basic approach was a non-experimental cross-sectional design which studies the relationship between the predictor and criterion variables. The predictor variables--synonymous with the independent variable in experimental designs--were used to predict the criterion. The criterion variable--synonymous with the dependent variable in experimental designs--is an outcome variable which may be predicted by one or more predictor variables. Hatcher (1994:10) notes that a weakness of non-experimental designs is that they are generally not able to support strong inferences of a causal nature. Linear regression is used to assess the predictiveness of the seven constructs: "Focus," "Agenda," "Facilitator," "Written Preparation," "Oral Presentation," "Team Characteristics," and "Reviewer Characteristics."

Assumptions

There are three underlying assumptions in this research. First, it is assumed that the Air Force environmental program criteria (refer back to Figure 8) are the most appropriate goals for peer review process effectiveness. Second, since these nine criteria are not being measured directly by the seven constructs, this research is based on the assumption that people's perceptions are strongly correlated to reality. Last, since this research focuses on improving the peer review process, the assumption is if the process is improved, the quality of the final product and services will also be improved (DoAF, 1991:3-3).

Sample and Setting

Hatcher (1994:12) defines a population as "the *entire collection* of a carefully defined set of people, objects, or events." Hatcher (1994:12) defines a sample as "a *subset* of the people, objects, or events selected from the population." Inferential statistics are used to generalize findings from a sample to some larger population (Hatcher, 1994:13). Figure 12 shows the populations and samples used in this research.

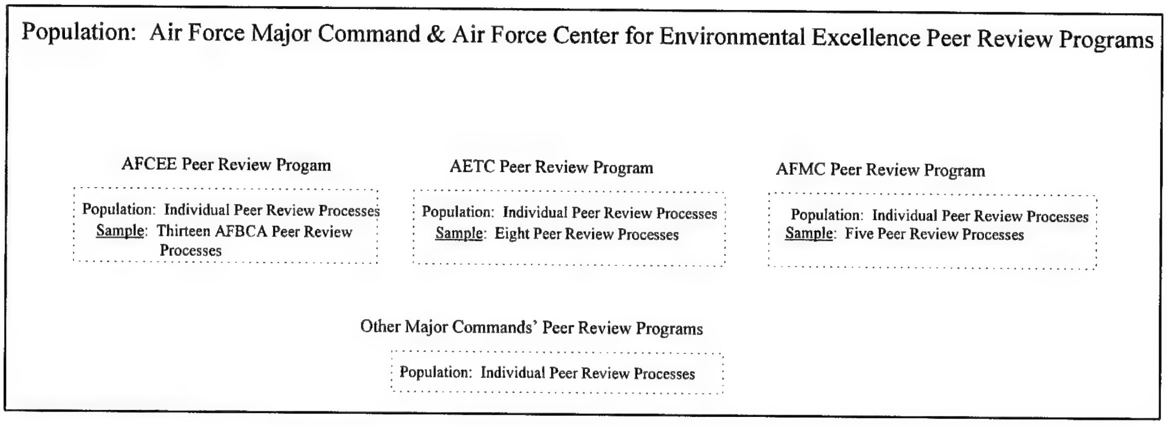

Population: Air Force Major Command & Air Force Center for Environmental Excellence Peer Review Programs
AFCEE Peer Review Progam
Population: Individual Peer Review Processes
Sample: Thirteen AFBCA Peer Review Processes
AETC Peer Review Program
Population: Individual Peer Review Processes
Sample: Eight Peer Review Processes
AFMC Peer Review Program
Population: Individual Peer Review Processes
Sample: Five Peer Review Processes
Other Major Commands' Peer Review Programs
Population: Individual Peer Review Processes

Figure 12. Populations and Samples

Several different populations were analyzed in this research. The most general population was the Air Force major commands and Air Force Center for Environmental Excellence (AFCEE) environmental peer review program; these were grouped together as a population because they were under the general policy guidance of the Air Force Center for Environmental Excellence (AFCEE, 1996:1). The peer review programs from the AFCEE, the Air Education and Training Command (AETC), and the Air Force Materiel Command (AFMC) were sampled to generalize to the Air Force major commands and AFCEE population.

Each of the peer review programs analyzed were composed of a population of individual peer review processes. Since the peer reviews were a dynamic process, these were not discrete populations. Further, the number of peer review processes composing the population changes from year to year and was different between commands and agencies. For the Air Force Base Conversion Agency program administered by AFCEE, two processes were observed and individuals participating in thirteen peer review processes were surveyed. For the Air Education and Training Command, individuals participating in eight processes were surveyed. For the Air Force Materiel Command, five processes were observed and individuals participating in five processes were surveyed.

Units of analysis are defined as the person, object, or event under study. Data collected on the units of analysis were aggregated to describe the sample and manipulated to make inferences about the population (Babbie, 1990:53-56). The units of analysis used to evaluate peer review process were individuals participating in the peer reviews. The units of analysis were aggregated and analyzed to draw inferences for the peer review processes.

Survey Variables

Peer Review Process Effectiveness The effectiveness construct is grounded in Air Force policy. The effectiveness construct was measured by twelve survey items with response alternatives ranging on a five-point Likert scale from (1) "strongly agree" to (5) "strongly disagree." The twelve items were additively combined to produce an overall effectiveness criterion measure.

The first ten items directly measure the degree to which the peer review process addresses the four-fold peer review program purpose (Figure 9 in the Survey Development Chapter). The fourth purpose was broken down into two subpurposes for a total of five areas to measure. These five areas were each measured from two perspectives for a total of ten items. They were measured from the perspective of the presentation from the base program manager and from the perspective of the peer review teams analysis of the presentation.

To illustrate the measurement of a program purpose, the first purpose was to evaluate the technical merits of the rationale used to set remediation goals (refer back to Figure 9). This purpose was measured from two perspectives by two survey items: "The Base Program Manager's presentation of proposed remedial actions was clear" and "The Peer Review Team completely reviewed the technical merits of all proposed remedial actions."

Two additional items were added to measure the perception of effectiveness of the person surveyed. A sample item states, "This Peer Review was worth the time and expense required to conduct it."

Peer Review Process Focus. The focus predictor variable was an additive combination of twelve items with response alternatives ranging on a five-point Likert scale from (1) "strongly agree" to (5) "strongly disagree." The twelve items were based on data from the related literature, Air Force policy, observation, and interviews. Items were chosen based on triangulation from several data sources. A sample item states, "The desired end-result of the Peer Review was stated with precision." This item, for example, was based on data from related literature (ACEC, 1990:3), the Peer Review Guidance Document (1996:6), observation, and interviews.

Agenda. The agenda construct is a nominal variable that measures whether or not an agenda was used in the peer review. The questionnaire instructions read, "If there was not a planned agenda for the peer review, please skip to question 28." If item 25 was answered, it is assumed there was an agenda used, and the agenda variable is assigned the value "1." If the question is skipped, the agenda variable is assigned the value "0."

Facilitator. The facilitator construct is a nominal variable that measures whether or not a facilitator was present in the peer review. The questionnaire instructions read, "If there was not a facilitator for the peer review, please skip to question 29." If item 28 was answered, it is assumed that there was a facilitator and the facilitator variable is assigned the value "1." If item 28 is skipped, facilitator is assigned the value "0."

Written Preparation. Written preparation was measured by an additive combination of four items with response alternatives ranging on a five-point Likert scale from (1) "strongly agree" to (5) "strongly disagree." A sample item states, "The Peer Review Team had adequate time to review documentation on projects under review provided by the Base Program Manager."

Oral Presentation. Oral presentation was measured by eight survey items. Six of these items have response alternatives ranging on a five-point Likert scale from (1) "strongly agree" to (5) "strongly disagree." A sample item states, "The Base Program Manager's briefing was at the right level of detail for efficient review." The remaining two items have response alternatives ranging on a five-point Likert scale from (1) "far too high" to (5) "far too low." A sample fill-in item states, "The level of detail of briefings presented by the Base Program Manager was" [response between (1) and (5)].

Peer Review Team Composition. The peer review team composition was measured by an additive combination of six items. The first two items have response alternatives ranging on a five-point Likert scale from (1) "far too high" to (5) "far too low." A sample fill-in item states, "The number of specialists on the Peer Review Team was" [response between (1) and (5)]. The last four items have response alternatives ranging on a five-point Likert scale from (1) "strongly agree" to (5) "strongly disagree." A sample item states, "The mix of specialists on the Peer Review team was right."

Peer Reviewer Characteristics. The peer reviewer characteristics construct was measured by an additive combination of six items having response alternatives ranging on a five-point Likert scale from (1) "strongly agree" to (5) "strongly disagree." A sample item states, “The Peer Reviewers were biased.” This item was the only item on the survey that was negatively scored. It was included to check response sets.

Survey Data Plan

Collection. The data collection plan was aimed at collecting data that will numerically support or contradict the seven constructs’ prediction of the criterion variable effectiveness. Since peer reviews are usually infrequent, the timeframe of this thesis does not allow many peer review processes to be captured. Under normal circumstances, the sample size from the population would be calculated based on the desired reliability or risk of the results. Because of the inherent difficulty in collecting data from the population, every attempt was made to maximize the sample size.

To facilitate tracking the surveys, each survey was numbered. This numbering served several purposes. First, the numbering allowed the data collector to track which surveys were not returned for follow-up. Second, numbering allowed the returned survey to be matched to a name for follow-up if survey items were inadvertently skipped. Third, numbering allowed the researcher the ability to contact survey respondent for additional information if qualitative responses on survey were unclear.

Reduction. For ease of statistical analysis, the data was handled by computer. The SAS® System computer program was used to calculate descriptive and inferential statistics from the raw data. The survey was designed to facilitate both completion and data entry; survey responses were all along the right margin of the page, except the last page with job information and qualitative open-ended items.

To facilitate data entry, quality control, and to minimize errors, periods were used to represent missing data. This allows the person performing data entry to ensure all responses were entered by graphically checking the width of the column. Further, columns of data can be checked against certain page locations using this global formatting methodology. Last by inspecting the location of periods, trends in missing data can be analyzed and corrected.

Follow-Up. Following a review of the initial data, trends were observed in missing data from surveys. To correct these problems, survey respondents with two or fewer missing data points were contacted by telephone to fill the data gaps. Six survey participants were contacted and data was obtained in each case. Filling the data gaps increased the usable sample size which increased the power of statistical analysis.

Statistical Analysis

Appendix H describes background information for the statistical methods used in this research.

Multi-Item Variables. Five of the survey predictor variable constructs--focus, written preparation, oral presentation, review characteristics, and team characteristics--are quasi-interval. Two of the survey predictor variable constructs--agenda and facilitator--are nominal.

Descriptive & Inferential Statistics. The SAS® System computer program was used to manipulate data and calculate both descriptive and inferential statistics (Appendix I). The SAS® System was chosen because it is a very powerful statistic tool. For descriptive statistics, two general procedures were used: "proc means" and "proc freq."

Descriptive Statistics. "Proc means" was used to display the mean, variance, minimum, maximum, and range of each variable for each survey. The variables described include both the item variables--one per survey item--and the multi-item variables. This data was useful to determine if the whole Likert scale was being used. Further, it was used to validate that certain items were reverse coded. Reverse coded items were items that have a positive response on the opposite end of the Likert scale than the "standard" item. Reverse coded items must be manipulated in the SAS® code before calculating any inferential statistics (Hatcher, 1994:102-104). Last, the "proc means" data was used to check for incorrectly entered data that could show as an "outlier." If, for example, the Likert scale uses responses from "1" to "5," a maximum value of "7" would indicate incorrectly entered data (Hatcher, 1994:447). While "proc means" is a basic procedure, the output can be quite powerful.

Frequency Distributions. "Proc freq" was used to display the frequency of each response to the Likert scale. This data was very useful in noting trends in responses. By calculating the percentages of each Likert scale response, the resulting frequency distribution depicts any trends (Appendix M). The frequency distributions were more useful than simply reporting the mean since the data can be mutimodal. Bimodal data, for example, could be clustered at responses "2" and "4" on a "1" to "5" scale. The mean of the data was "3," although "3" isn't necessarily used by any respondent (Hatcher, 1994:105-107).

Test for Multicollinearity. For inferential statistics, several SAS® procedures were used to build the linear regression model, and to test its parameters. "Proc corr" was used to test for correlation between the predictor multi-item variables. This type of correlation, called multicollinearity, can cause problems with the regression model such as: bias, incorrect sign, or failure to show statistical significance (Hatcher, 1994:447).

Test for Internal Consistency. "Proc corr" was also used to test for internal consistency between the items that compose the multi-item variables. By adding "alpha" to the options in the correlation procedure requests, coefficient alpha be computed for the group of variables included in the "var" statement. "A widely used rule of thumb of .70 has been suggest," as a minimum acceptable coefficient alpha (Hatcher, 1994:513). However, Hatcher (1994:513) also noted that coefficient alphas below .70 were often accepted in social science literature. SAS® reports two values for "alpha;" the "alpha" value for "RAW variables" is the value usually reported (Hatcher, 1994:512).

Hatcher (1994:513) also points out that it is possible to increase the reliability of a multi-item scale. By reviewing the correlations generated by "proc corr" for each item composing the scale, the researcher looks for extremely low or negative correlations. If this occurs, either the item was reverse coded, or it may be measuring another construct. Further, it is possible to add items to increase the internal consistency.

Linear Regression Model. SAS® procedure "proc reg" with option "r-square" was used to generate the R^2 for each combination of variables used in the linear regression model. These R^2 values were equal to those computed using the "forward" option to generate the model. The difference is, the "r-square" option prints the R^2 for all combinations of predictor variables, not just the ones that were statistically significant. This printout was useful to look for trends in variable's ability to account for variance in the criterion.

"Proc reg" was used next to build the linear regression model. In the linear regression model, variables were rotated into the linear regression equation based on a selected criteria. The model was generated using the option "forward" which builds the model in a sequential process based on the variable's R^2 value. The R^2 value indicates the percent of variance that the predictor variable accounts for in the criterion (Hatcher, 1994:429). The procedure continues to add the variable with the highest R^2 into the model until no variable meets the .5 tolerance for entry into the SAS® model. (Note: the level of significance for this study is .05; the tolerance for entry into the SAS® model was .5.)

While variables were entered into the model based on R^2, they must also be statistically significant. SAS® prints both the F-value and its associated P-value from the test of significance. Variables with P-values greater than .5 did not meet the tolerance level for the SAS® model. Statistically insignificant means we cannot conclude that the

R^2 value was greater than zero in the population. A 95% level of significance was used in this study, accordingly variables with P-values greater than .05 were rejected.

Confidence Level. Choosing the right confidence level for a study was not a simple decision. For confidence intervals, more is not necessarily better. While a 99% confidence level is more reliable than a 95% interval, it also has lower precision. Confidence intervals are inversely related to precision; as confidence intervals increase, precision decreases. There are three confidence levels that appear frequently in the literature: 90 percent, 95 percent, and 99% (Devore, 1995:280-281). A 95% confidence level was used in this study to balance reliability with precision.

Uniqueness Index. The uniqueness index was computed to determine how much variance each predictor variable accounts for in the linear regression equation. If more than one variable has an R^2 that was statistically significant, then the uniqueness index was used to determine how much additional variance the entered variable adds to the model. The equation is $U = R^2_{\text{Full}} - R^2_{\text{Reduced}}$, where "U" is the uniqueness index, R^2_{Full} is the R^2 from the equation including the newly entered variable, and R^2_{Reduced} is the R^2 from the linear regression equation before the new variable is entered (Hatcher, 1994:432).

Chapter 5

Results

General Survey Responses

A total of 141 surveys were administered, with a 50% response rate (N = 70). The sample size used in the analysis was reduced due to missing data. Sample size is noted for each analysis. Survey data from the following organizations is contained in the general sample: the Air Force Center for Environmental Excellence, the Air Education and Training Command, and the Air Force Materiel Command.

Air Force Center for Environmental Excellence. The Air Force Center for Environmental Excellence held thirteen peer reviews for the Air Force Base Conversion Agency from 27 February 1996 through 6 March 1996. Each of these peer reviews covered the environmental projects at a particular closure base. A core of six to eight peer reviewers was scheduled for each panel. Many of these reviewers participated in more than one peer review. On 12 April 1996, 112 surveys were mailed to people who participated in the environmental project peer reviews. The responses were requested by 30 June 1996. Surveys were numbered to facilitate follow-up. The response rate was 41% (N=46).

Air Education and Training Command. The Air Education and Training Command conducted an environmental remediation project peer review from 19 April 1996 through

10 May 1996. Eight bases presented projects for peer review. Surveys were mailed under letter from HQ Air Education and Training Command to each base Remedial Program Manager on 1 July 1996. The responses were requested by 18 July 1996. Eight surveys were mailed and returned completed for 100% response rate (N=8).

Air Force Materiel Command. The Air Force Materiel Command conducted an environmental remediation project peer review from 20 August 1996 through 22 August 1996. Five bases presented projects for review. Surveys were administered at the end of each peer review session. A total of 21 surveys were administered to peer review participants; the response rate was 76% (N=16).

Demographics

Figure 13 shows the distribution of the 68 surveys returned; two surveys were returned incomplete. The majority of the survey respondents were from the Air Force Center for Environmental Excellence. These respondents served in a variety of capacities ranging

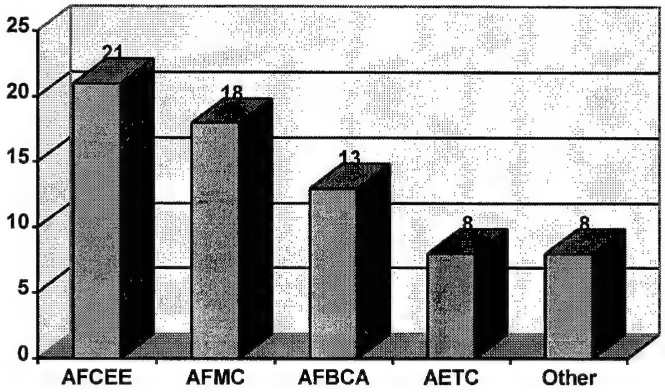


Figure 13. Survey Respondents' Assigned Organizations

form managerial oversight to technology transfer. The "Other" category of respondents includes sources such as the Environmental Protection Agency laboratory.

Figure 14 shows the breakout of the job positions held by the survey respondents. The largest group of respondents were the project managers who managed the environmental remediation projects. The next largest group of respondents were from non-categorized jobs such as lawyers, regulators, and observers. Sixteen respondents were either engineers or scientists on the peer review team. Eight were present for management oversight. Two respondents were from the Technical Transfer section at the Air Force Center for Environmental Excellence.

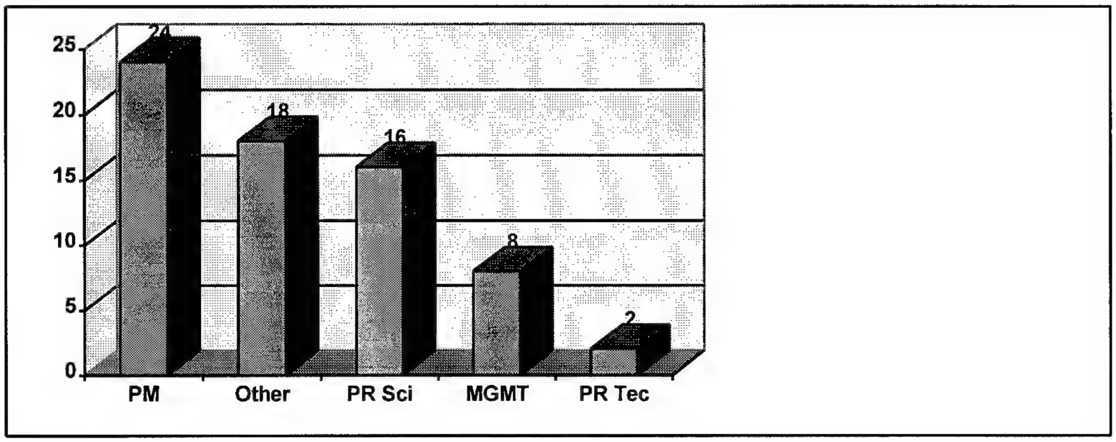


Figure 14. Survey Respondents' Job Positions

Peer Review Effectiveness

Individuals were asked to respond to twelve items pertaining to peer review effectiveness. The first item asked about the overall effectiveness of the process. The second item asked whether the peer review was a good use of Air Force resources. When asked to respond to, "This overall peer review process was effective," most respondents agreed that the overall peer review was effective (Figure 15). When asked to respond to, "This peer review was worth the time and expense required to conduct it," there was a bimodal frequency distribution (Figure 16). Survey responses were heavily weighted between strongly agree and agree or at disagree. While most respondents felt the peer review process was effective, some felt that the use of resources was not worthwhile. This possibly indicates people felt that too many resources were used for peer reviews. This finding is further supported by open-ended questions were people indicated that there were often too many reviewers in general or too many reviewers from a specific discipline.

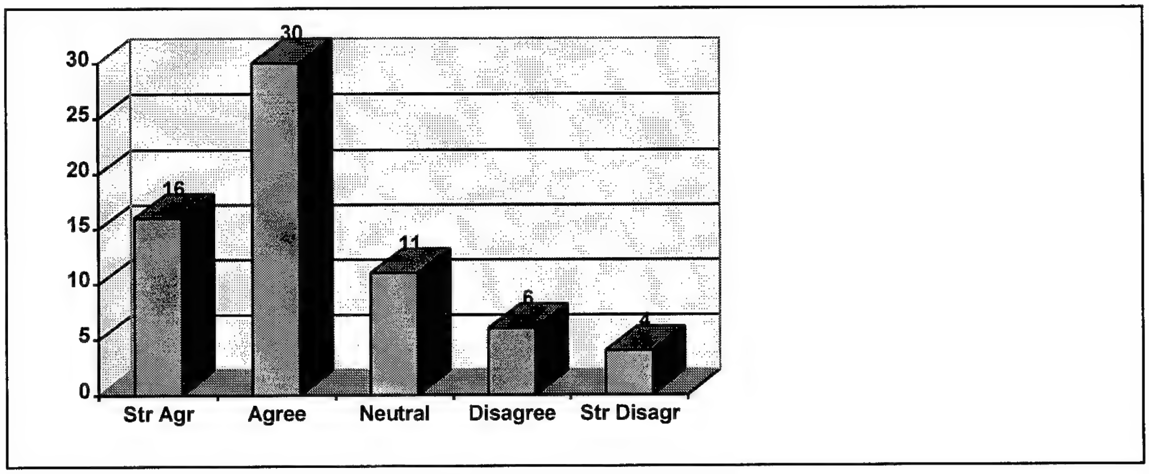


Figure 15. Overall Peer Review Effectiveness Responses

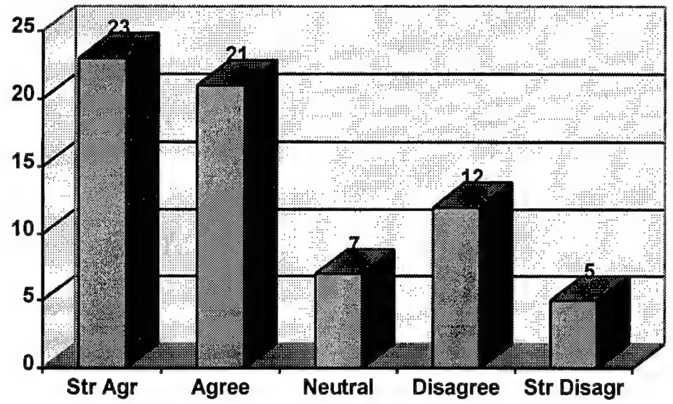


Figure 16. Utilization of Resources Responses

Descriptive Statistics and Correlations

Results were analyzed using descriptive statistics and correlation analysis. Means, standard deviations, item-total correlations, and Cronbach alpha values appear in Tables 2 and 3. Mean values are, for the most part, between 2.0 and 3.0 indicating respondents assessed the positively-worded items between "agree" and "neutral." Standard deviation values range from 0.60 to 1.30. Item-total correlations were all positive, except where noted. Cronbach alpha values range from 0.60, for peer review team characteristics to 0.91, for oral presentation. Nunnally (1978) recommends Cronbach alpha values of .70 or higher. Mean values and standard deviations for survey items and composite variables were also summarized in Tables 2 and 3.

TABLE 2.
CRITERION MEASURE ITEM-TOTAL CORRELATIONS & SCALE ALPHA[a]

Criterion Variable	Mean[b]	Standard Deviation	Item-Total Correlation	Cronbach Alpha
Effectiveness of the Peer Review Process	**2.33**	**0.68**		**0.90**
Detailed funding estimates[c].	2.34	1.07	0.65	
Clear proposal of remedial actions	2.09	0.85	0.59	
Alternate remedial actions were clear.	2.51	0.99	0.57	
Clear rationale to meet regulatory requirements.	2.04	0.87	0.54	
Risk based approach was clear.	2.41	0.97	0.44	
Complete review of the technical merits of all proposed actions.	2.24	0.89	0.70	
Complete review of the technical merits of all alternatives.	2.60	1.01	0.69	
Complete review of the PM's[d] rationale to meet regulations.	2.16	0.95	0.74	
Completely review of the PM's risk assessment.	2.65	0.96	0.64	
Adequately review of funding estimates.	2.35	0.96	0.49	
Effective peer review process..	2.28	1.13	0.73	
Peer review was worth the time and expense.	2.34	1.30	0.59	

[a] N = 65, Values in bold type represent values from the composite variable.

[b] Rating scale: 1 = strongly agree; 2 = agree; 3 = neutral; 4 = disagree; and 5 = strongly disagree. [c] For exact item wording see Appendix E. [d] Project Manager's.

Composite Variables Correlations

Results were analyzed using both bivariate correlation and multiple linear regression. A correlation matrix was first generated to determine the relationships between composite variables (Table 4). Oral presentation has the strongest relationship ($r = 0.74$, $p < .0001$) with effectiveness. Effectiveness was also significantly correlated with focus ($r = 0.65$, $p < .0001$), written preparation ($r = 0.65$, $p < .0001$), reviewer characteristics ($r = .47$, $p < .0001$), and team characteristics ($r = 0.40$, $p < .002$). Agenda and facilitator were not significantly correlated with effectiveness. All variables had a positive correlation with effectiveness, as expected.

TABLE 3.
PREDICTOR VARIABLE ITEM-TOTAL CORRELATIONS & SCALE ALPHAS

Predictor Variables	Mean[a]	Standard Deviation	Item-Total Correlation	Cronbach Alpha
Focus[b]	**2.47**	**0.60**		**0.87**
Precisely stated end-result of the peer review.	2.36	0.95	0.61	
Well defined scope.	2.39	0.95	0.52	
Focused on validating funding.	2.59	0.95	0.44	
Focused on cost savings.	2.28	0.95	0.43	
Focused on evaluating PMs rational to meet regulations.	2.23	0.83	0.41	
Focused on validating assumptions.	2.25	0.78	0.62	
Focused on validating designs.	2.63	1.06	0.56	
Focused on reviewing all non-standard assumptions.	2.70	0.95	0.67	
Focused on non-standard designs.	2.97	0.89	0.63	
Focused on scientific and technical aspects of projects.	2.06	0.96	0.60	
Focused on evaluating alternative technologies.	2.48	0.96	0.60	
Focused on evaluating risk assessment results.	2.72	0.92	0.66	
Written Preparation[e]	**2.64**	**0.84**		**0.79**
PRT[f] had adequate time to review documentation.	2.79	1.12	0.53	
PRT read documentation.	2.95	1.16	0.54	
Written material was well organized.	2.33	1.03	0.67	
Written material was at the right level of detail.	2.49	1.01	0.65	
Oral Presentation[g]	**2.44**	**0.71**		**0.91**
Briefing at right level of detail.	2.25	1.00	0.82	
Detailed project data readily available.	2.41	1.10	0.79	
Project locations clearly described.	2.01	0.90	0.74	
Clear description of overall remediation program.	2.10	0.93	0.74	
Clear description of individual projects.	1.99	0.85	0.82	
PM sold project to PRT.	2.59	1.10	0.66	
Project documents at right level of detail.	3.18[h]	0.72	0.47	
Briefing at right level of detail.	3.09[h]	0.62	0.61	
Peer Review Team Characteristics[i]	**2.54**	**0.44**		**0.60**
Number of specialists on PRT.	2.88[h]	0.62	-0.15[j]	
Experience of specialists on PRT.	2.91[h]	0.55	0.32	
PRT was objective.	2.18	0.82	0.54	
PRT was constructive.	2.13	0.76	0.56	
PRT adequately considered community concerns.	2.76	0.95	0.48	
Mix of specialists on PRT was right.	2.39	0.82	0.27	
Peer Reviewer Characteristics[k]	**2.32**	**0.64**		**0.78**
Reviewers were unbiased[l].	2.65	1.08	0.51	
Reviewers were independent of projects under review.	2.16	0.94	0.45	
Reviewers had adequate knowledge.	2.47	1.04	0.43	
Reviewers were good listeners.	2.18	0.85	0.72	
Reviewers were good interviewers.	2.24	0.81	0.53	
Reviewers were good communicators.	2.22	0.81	0.61	

[a] Rating scale: 1 = strongly agree; 2 = agree; 3 = neutral; 4 = disagree; and 5 = strongly disagree. [b] N = 64, Values in bold type represent values from the composite variable. [d] Project Manager. [e] N = 66. [f] Peer Review Team. [g] N = 67.
[h] Rating scale: 1 = far too high; 2 = somewhat high; 3 = ok; 4 = somewhat low; and 5 = far too low. [i] N = 67. [j] Due to the negative internal consistency for the first item in the composite Peer Review Team Characteristics predictor variable, the usefulness of this item in the model was reviewed. By removing the item from the composite variable (Hatcher, 1994:513), the Cronbach Alpha was raised to 0.70; however, the linear regression model was weakened. When the "Number of specialists on PRT" item was deleted from the linear regression model, the total R^2 was decreased from 0.73 to 0.72, a small yet significant loss of predicted variance. A negative correlation can also indicate reverse coding. When "Number of specialists on PRT" was reverse coded, the total R^2 was again decreased from 0.73 to 0.72; therefore, this item was retained in the model. [k] N = 68. [l] Item was reverse coded in survey.

This analysis of correlation indicates that effective peer reviews are strongly correlated to the quality of the oral presentation. For example, if the oral presentation is inadequate, it is likely that the peer review is not effective. Likewise, if the peer review is focused, it is likely to be effective. The correlation matrix is one indication of the predictiveness of the constructs. Linear regression is next used to further assess the correlation.

TABLE 4.
CORRELATION MATRIX

	Intercorrelations							
Variable	Effectiveness	Focus	Agenda	Facilitator	Written	Oral	Team	Reviewer
Effectiveness	-							
Focus	**0.65***	-						
Agenda	0.04	-0.01	-					
Facilitator	0.19	0.14	**0.47***	-				
Written Preparation	**0.65***	**0.47***	0.03	0.21	-			
Oral Presentation	**0.74***	**0.43***	0.02	0.13	**0.66***	-		
Team Characteristics.	**0.40***	**0.70***	-0.20	-0.00	0.24	0.24	-	
Reviewer Char.	**0.47***	**0.69***	-0.06	0.14	0.20	0.16	**0.74***	-

[a]N = 61. *p<.05

General Linear Regression Model Results.

The general linear regression model was used to assess the predictive relationship between the seven composite variables and the criterion, peer review effectiveness. Unstandardized regression coefficients, beta weights, uniqueness indices (partial R^2), and p values appear in Table 5. The beta weights (standardized regression coefficients) and the uniqueness indices were used to assess the relative importance of the variables. Oral presentation produced the largest beta weight at 0.50 ($p < .0001$) followed by reviewer characteristics with a beta weight of 0.29 ($p < .01$) and focus with a beta weight of 0.27 ($p < .03$). All significant beta weights were in the predicted direction. Oral presentation and focus predicted 68% of the variance in peer review effectiveness. Oral presentation contributed 0.54 to the R^2 value, while focus contributed 0.14.

Air Force Center for Environmental Excellence. Stepwise linear regression was used to assess the predictive ability of the seven composite variables. Unstandardized regression coefficients, beta weights, uniqueness indices, and p values appear in Table 6 for the reduced AFCEE sample. The beta weights and the uniqueness indices were used to assess the relative importance of the variables. Oral presentation produced the largest beta weight at 0.49 ($p < .0007$) followed by reviewer characteristics with a beta weight of 0.35 ($p < .05$). All significant beta weights were in the predicted direction.

The stepwise linear regression model was run with a level of significance of 0.05. Oral presentation and reviewer characteristics predict 62% of the variance in peer review effectiveness. Oral presentation contributes 0.49 to the R^2 value, while reviewer characteristics contributes 0.13. The findings regarding the beta weights matched those for uniqueness indices.

Air Education and Training Command. Stepwise linear regression was used to assess the predictive ability of the seven composite variables. Unstandardized regression coefficients, beta weights, uniqueness indices, and p values appear in Table 7 for the reduced AETC sample. The beta weights and the uniqueness indices were used to assess the relative importance of the variables. Focus produced the largest beta weight at 0.84

TABLE 5. RESULTS OF STEPWISE LINEAR REGRESSION[a] (General)

	Dependent Variable: Peer Review Effectiveness		
Independent Variables[b]	Standardized Regression Coefficient (BETA)	Uniqueness Index (Partial R^2)[c]	F Test
Oral Presentation	**0.50***	**0.54***	$p < .0001$
Focus	**0.27***	**0.14***	$p < .0001$
Reviewer Characteristics	**0.29***	0.02	$p < .07$
Written Preparation	0.17	0.02	$p < .09$
Team Characteristics	-0.17	0.01	$p < .14$

[a] N = 61. [b] Predictor variables agenda and facilitator did not meet the 0.50 significance level for inclusion into the model. [c] R^2 for all five variables was 0.73. *$p<.05$

TABLE 6. RESULTS OF STEPWISE LINEAR REGRESSION[a] (AFCEE)

	Dependent Variable: Peer Review Effectiveness		
Independent Variables[b]	Standardized Regression Coefficient (BETA)	Uniqueness Index (Partial R^2)[c]	F Test
Oral Presentation	**0.49***	**0.49***	$p < .0001$
Reviewer Characteristics	**0.35***	**0.13***	$p < .0008$
Written Preparation	0.20	0.03	$p < .11$
Facilitator	0.08	0.02	$p < .20$
Focus	0.20	0.01	$p < .39$
Team Characteristics	-0.17	0.01	$p < .31$

[a] N = 42. [b] Predictor variable agenda did not meet the 0.50 significance level for inclusion into the model. [c] R^2 for all six variables was 0.68. *$p<.05$

TABLE 7. RESULTS OF STEPWISE LINEAR REGRESSION[a] (AETC)

	Dependent Variable: Peer Review Effectiveness		
Independent Variables[b]	Standardized Regression Coefficient (BETA)	Uniqueness Index (Partial R^2)[c]	F Test
Focus	**0.84***	**0.69***	$p < .03$
Oral Presentation	**0.52***	**0.26***	$p < .02$
Written Preparation	0.15	**0.04***	$p < .04$
Agenda	-0.12	0.00	$p < .33$
Facilitator	0.13	0.00	$p < .13$

[a] N = 7. [b] Predictor variables team characteristics and reviewer characteristics did not meet the 0.50 significance level for inclusion into the model. [c] R^2 for all five variables was 1.00. *$p<.05$

TABLE 8. RESULTS OF STEPWISE LINEAR REGRESSION[a] (AFMC)

	Dependent Variable: Peer Review Effectiveness		
Independent Variables[b]	Standardized Regression Coefficient (BETA)	Uniqueness Index (Partial R^2)[c]	F Test
Oral Presentation	0.27	**0.71***	$p < .0006$
Focus	0.41	**0.11***	$p < .05$
Facilitator	**0.33***	**0.07***	$p < .05$
Written Preparation	0.14	0.04	$p < .10$
Reviewer Characteristics	-0.44	0.01	$p < .32$

[a] N = 12. [b] Predictor variables agenda and team characteristics did not meet the 0.50 significance level for inclusion into the model. [c] R^2 for all five variables was 0.94. *$p<.05$

($p < .02$) followed by oral presentation with a beta weight of 0.52 ($p < .03$), and written preparation with a beta weight of 0.15 ($p < .09$). All significant beta weights were in the predicted direction.

The stepwise linear regression model was run with a level of significance of 0.05. focus, oral presentation, and written preparation predict 99% of the variance in peer review effectiveness. Focus contributes 0.69 to the R^2 value, while oral presentation contributes 0.26 and written preparation contributes 0.04. The findings regarding the beta weights matched those for uniqueness indices.

Air Force Materiel Command. Stepwise linear regression was used to assess the predictive ability of the seven composite variables. Unstandardized regression coefficients, beta weights, uniqueness indices, and p values appear in Table 8 for the reduced AFMC sample. The beta weights and the uniqueness indices were used to assess the relative importance of the variables. Facilitator presented the only significant beta weight at 0.33 ($p < .02$) using a .05 level of significance. This beta weight was in the predicted direction.

The stepwise linear regression model was run with a level of significance of 0.05. oral presentation, focus, and facilitator predict 89% of the variance in peer review effectiveness. Oral presentation contributes 0.71 to the R^2 value, while focus contributes 0.11 and facilitator contributes 0.07.

The findings regarding the beta weights do not matched those for uniqueness indices. The only significant beta weight was facilitator ($\beta = .33$, $p < .02$); however, both oral presentation and focus contribute more to the R^2 value than facilitator.

Open-Ended Items

There were four fill-in-the-blank items at the end of the survey. Of the four, the following two questions are pertinent to present as results:

1. The following type of specialist should have been on the Peer Review Team:
2. The Peer Review Process would be improved by...

The responses to these items are transcribed word-by-word in the appendices. Appendix I contains the comments from the Air Force Base Conversion Agency surveys. Appendix J contains the comments from the Air Education and Training Command surveys. Appendix K contains the comments from the Air Force Materiel Command surveys.

Air Force Center for Environmental Excellence. Five out of thirteen respondents suggested adding a cost analyst to the peer review team. Thirteen out of thirty-four respondents to the second question noted that preparation is a weakness. Documents were not given to reviewers with adequate time for review and reviewers did not thoroughly read the documents prior to the meeting.

Air Education and Training Command. The Air Education and Training Command survey comments did not significantly recommend any change in the mix of specialists.

Two out of the seven responses suggested more concise guidance, a prepublished agenda, and typed minutes following the peer review.

Air Force Materiel Command. The Air Force Materiel Command survey comments did not significantly recommend any change in the mix of specialists. Five out of thirteen responses recommended the peer review objectives be clearly outlined prior to the review. In addition, three responses suggested the reviewers send specific questions to the bases requesting clarification on unclear aspects of the projects. Two of the responses recommended the peer review leader send written recommendations from the panel to the bases.

Observation Results

Both Air Force Center for Environmental Excellence and Air Force Materiel Command peer reviews were observed. The Air Education and Training Command data was gathered by survey, AETC peer reviews were not observed. The Air Force Center for Environmental Excellence administered the peer review for the Air Force Base Conversion Agency. Appendix B contains the observations from the Air Force Base Conversion Agency peer reviews. The AFCEE administered peer review observations were used both to develop the survey instrument and to further assess the survey results and develop recommendations. Appendix L contains the observations from the Air Force Materiel Command peer reviews; the AFMC observations were used to qualitatively assess the survey results and develop recommendations.

Air Force Center for Environmental Excellence. Observation of AFCEE administered peer reviews was limited to a closed base in California and a base on the closure list in Texas. One problem present in both reviews--and from interviews is common in most reviews--is the peer review team did not adequately review the written documentation. Several reasons were given for this deficiency: the documentation was not, in some cases, given to AFCEE with sufficient lead time; AFCEE, in-turn, did not give the documentation to the reviewers with sufficient lead time; and the reviewers were not allocated sufficient time to accomplish the review.

Further, the written documentation was very lengthy and not well organized. Two problems were cited by peer review participants for this problem. The first is a lack of clear and concise guidance from the AFCEE and the AFBCA. Second, the bases did not seek feedback from AFCEE while developing the written documentation.

The presentations were weak in several areas. First, overhead projectors were not used to give an overview of the site. This was a handicap for clearly explaining the projects under review. Also, the project managers did not try to "sell" the projects. The attitude of some was, "why do I have to be here?" A more clearly defined scope and focus coupled with additional motivation and guidance could help with these problems.

Air Force Materiel Command. The discussions and observations of the Air Force Materiel Command peer reviews provided additional insights. A common area of confusion was the importance of non-technical issues in the peer review process. In most of the reviews, non-technical issues--such as political, public, and regulatory concerns--

were given as partial justification for decisions. In the past, the AFMC peer review process was used as a pre-approval authority and weighed all justifications including non-technical issues into the final decision of project appropriateness. Many project managers were confused because the scope of the peer review was recently limited to primarily technical issues.

The presentation was weak in several of the reviews. In reviews that used the overhead and gave an overview of the projects under review, there was less observed reviewer confusion. The Hill AFB peer review utilized both an overhead and an overview; the result was a clear and concise presentation.

One weakness of the reviews was a lack of clear guidelines. There was an informal facilitator whose role was often confusing. An agenda would have focused the review and allowed the project managers to prepare for review team questions. Further, misunderstandings developed, in some cases, because project managers were not clear exactly what was expected of them.

Chapter 6

Discussion

Introduction

The survey results are assessed in light of the interviews and observations. The survey results are confirmed by both interviews and observations. The results indicate that the peer review process is currently effective and effectiveness can be improved by measuring oral presentation and focus. Both oral presentation and focus were consistently mentioned as important peer review factors in interviews and were noticeably important during observations.

Effective Peer Review Process

Because the projects under peer review are few in number but often very complex, the current peer review strategy is aimed at effectiveness over efficiency. Of the reviews observed, the peer reviews often erred on the side of having too many experts. In the Air Force Center for Environmental Excellence peer reviews, for example, there was often a duplication of specialists. This is prudent in light of the immense value of the projects in comparison to the peer reviewer's salary.

The results in this study indicate that participants evaluated the current peer review processes as being relatively effective based on Air Force policy. Approximately 69% of those surveyed either agreed or strongly agreed that the overall peer review process is effective. Only 15% disagreed or strongly disagreed.

The results in this study indicate mixed support for utilization of resources on peer reviews. While 65% either strongly agreed or agreed, 25% either disagreed or strongly disagreed. The significant percentage of those who disagreed is expected given the prudent tendency to overload the peer reviews with technical experts.

General Model

For the general model, the results in this study indicate that oral presentation and focus are the statistically significant predictors of effectiveness. These results are supported by peer review process observations (Appendices B and L). Oral presentation accounted for 54% of the total variance. Focus accounted for 14% of the variance. These results indicate that if the Air Force improves the quality of oral presentation and focus, the peer reviews are likely to become more effective with respect to the requirements established by Air Force policy.

Oral Presentation. From peer review observation, oral presentations detracted from the effectiveness of the peer reviews. The use of overheads, in many cases, would have made the presentation much more effective. The positive impact of overheads is inferred from the confusion that was evident when a pictorial overview was lacking from an observed peer review. Further, the presenters did not give a verbal overview of the project(s) under review in many cases. An overview could have minimized misunderstandings of assumptions and designs that followed.

Focus. Focus was also a weakness in the observed peer reviews. None of the peer reviews observed had a published agenda to keep the review on target. Further, in none of the reviews did the reviewers give the project managers questions in advance in areas that were not clear from written documentation. In several cases, project managers were not clear what was expected from them by the peer review team. Survey respondents recommended:

1. more concisely and clearly communicated guidance;
2. an agenda distributed *before* the review; and
3. reviewers read documentation and send questions to project managers *before* the review.

Significant Variables by Major Command

The results in this study indicate the peer review constructs have different degrees of importance among major commands.

Oral Presentation. Oral presentation is either the first or second most important variable to predicting effectiveness in all cases. Based on the results, major commands should focus efforts on improving the oral presentation of projects at peer reviews. Oral presentation can be improved by encouraging the project managers to "sell" the projects through a clear description of the overall program utilizing visual aids such as photographs and overhead transparencies. Further, the project manager should have pertinent project information to answer specific questions posed by the peer review team. Ideally, these questions would be furnished to the project manager prior to the review.

Reviewer Characteristics. Reviewer characteristics was only statistically significant for the Air Force Center for Environmental Excellence (AFCEE) administered peer reviews. This indicates that for AFCEE peer reviews the peer reviewer's characteristics is a significant predictor of the effectiveness of the review. Review characteristics includes the degree to which reviewers are: good listeners, good interviewers, and good communicators. Further, reviewer characteristics also captures the perception of those surveyed of whether the reviewers were biased, independent of the project(s) under review, and had adequate knowledge to perform a thorough review. The results indicate that survey respondents perceived that the reviewers were independent of the projects under review, that they had adequate knowledge to perform the review, and that they were unbiased (See Table 3).

Focus. The results indicate that both the Air Education and Training Command (AETC) and the Air Force Materiel Command (AFMC) survey respondents perceived that peer review focus was important. For AETC, focus contributed 69% to the variance of peer review effectiveness. For AFMC, focus contributed 11% to the variance of effectiveness. The focus multi-item variable captures: whether the desired end-result of the peer review was stated with precision; whether the scope was well defined; and whether the peer review cover the areas as specified by Air Force policy.

Written Preparation. The results indicate that written preparation accounts for 4% of the variance of effectiveness for AETC. While written preparation is statistically significant for AETC, it accounts for a small portion of the variance as compared to focus and oral presentation. Resources should be focused on improving oral presentation and focus before they are focused on written preparation. Written preparation can be improved by focusing on and ensuring: the level of detail is adequate for written material; the reviewers receive the written material in advance of the peer review; and the reviewers read the written material prior to the review.

Facilitator. Similarly, the results indicate that facilitator accounts for 7% of the variance of effectiveness for AFMC. While facilitator is statistically significant for AFMC, it accounts for a small portion of the variance as compared to oral presentation. Accordingly, oral presentation is a more significant predictor than facilitator of peer review effectiveness. Improving the oral presentation is more likely to improve peer review effectiveness than would improvements in the facilitator construct.

The facilitator construct measures the perception of whether or not there was a facilitator. Based on observation, there was a team leader in each of the peer reviews; however, there was not a facilitator. Since the team leader has a different "agenda" than a formal facilitator, the facilitator role was not clearly filled. Guidance that outlines the roles of the participants' roles in the peer review process along with clearly communicating what role each participant fills, would help correct the misperceptions peer review participants encounter.

Peer Review Process Recommendations

Based on interviews (Appendix C), observations (Appendices B and L), and survey comments (Appendices I, J, and K) the recommendations in Figure 17 could improve the peer review processes. The Air Force Center for Environmental Excellence is developing a standard peer review process protocol for the Air Force, Peer Review Guidance Document (1996). This guidance document serves as a basis for continuous improvement and communication of a peer review protocol incorporating lessons learned for major command use.

Scope and Objectives. Professional organizations from the American Consulting Engineers Council (1990a) to the American Society of Civil Engineers (1988) stress the importance of a clearly defined scope and objectives. The Peer Review Guidance Document (AFCEE, 1996) proposes a standardized four-fold purpose. Observations have shown that peer reviews vary widely depending on the projects being reviewed and on the degree of development of the project's design. Survey responses and interviews further indicate that peer reviews vary widely; to respond to this variance, the scope and objectives of peer reviews should be tailored from the standardized four-fold purpose.

Peer Reviewer Mix. The American Society of Civil Engineer (1988:83) determines the size of the peer review team and the mix of specialists from the tailored scope and objectives. The Peer Review Guidance Document (AFCEE, 1996:2) recommends the peer review team expertise be drawn from areas including, "chemistry, hydrology, geology, environmental engineering, risk assessment, remediation technologies, project management, cost estimation, environmental law and environmental regulation."

One problem illuminated by interviews and open-ended survey comments was there are often duplicity of specialists. While having a risk assessor on the peer review team may be helpful, the decision-maker must also determine whether having three risk assessors is prudent.

Trained Facilitator. The two major commands observed utilized different methods to guide the peer review process. The Air Force Center for Environmental Excellence utilized a trained facilitator. This facilitator held a separate role from the team chief and was primarily concerned with ensuring the peer review process was effective in-so-much as it met the four-fold purpose outlined in Air Force policy. The Air Force Materiel Command, on-the-other-hand, utilized a dual role team leader/facilitator.

From observation, the separate facilitator role appears superior to the combined team leader/facilitator role because it allows the facilitator to focus on peer review effectiveness while the team leader role is often concerned with peer review efficiency. The team leader is concerned with reviewing all scheduled projects, utilizing all time efficiently to keep costs down, and streamlining the process. A facilitator is concerned with ensuring all projects are reviewed effectively. Combining the role leads to confusion for those being reviewed because of the subtle conflict of roles.

Prepublished Agenda. The Peer Review Guidance Document (AFCEE, 1996:6) recommends an agenda be published prior to the peer review. The American Consulting Engineers Council (1990a) recommends a minimum of a checklist, of items to be reviewed, be published before the review to facilitate preparation. None of the peer reviews observed had a prepublished agenda.

One weakness in all the observed peer reviews was a lack of focus. The facilitator or team leader was forced to establish the focus without advanced warning to those under review to facilitate preparation. A prepublished agenda would begin to correct this deficiency.

Visual and Verbal Overview. One area that was commonly critiqued on the peer review process was the lack of a clear and concise overview. Of the peer reviews observed, the presentations that utilized both visual and verbal overviews appeared more productive and focused. In reviews lacking an overview, time was wasted going back and forth between detailed questions and answers and giving general information.

Prioritization of Peer Review Goals. The prioritization in Figure 17, Recommendation 7, emerged from Air Force policy, observations, and interviews. Air Force policy requires that remediation projects are "risk-based, technically sound and cost appropriate" (AFCEE, 1996:5). Because the measure of merit for Air Force Base Conversion Agency peer reviews was cost savings, their priority was to reduce costs. This focus on reducing costs led the review toward costs discussions prematurely in the observed reviews.

Further, there was often confusion when the primary objective was stated to ensure the projects were technically sound, when the actual primary objective was to ensure the project were cost effective. Interviews and discussions with peer review team members further emphasized the importance of ensuring the need for the project first. If the risk does not justify action, it makes little sense to evaluate the technology for cost savings, unless the project is mandated for reasons other than risk such as by political pressure.

Additionally, it is logical that the first priority should be to ensure the project is really needed. "Does the risk justify remediation actions?" Once remediation is justified based on risk, the available technologies should be explored for the impacted media based on the risk to human health and the environment. The last step is to ensure the chosen remediation is cost-effective and that all costs are minimized. Choosing the correct technology is more important than cost because an expensive remediation project that lowers the risk to human health and the environment is arguably superior to a cost-effective project that does not lower risk to safe levels. However, this prioritization must serve only as a guide, maximum flexibility is essential to ensure the peer review experts and the decision-maker can tailor the decision to the specific conditions.

Peer Review Team as Technical Advisory Board. This recommendations is based primarily on interviews and observations of Air Force peer reviews. Many peer reviewers voiced frustration that they are being asked to evaluate projects technically and "politically" concurrently. "Politically" refers to the other issues that Air Force managers take into account when making decisions such as stakeholder concerns, regulatory body preferences, and political issues. Scientists and engineers are best suited to make technically based decisions; mangers are better informed and trained to make politically based decisions. Cozens (1987:80) said it well when she stated, "Scientists are intelligent, and they are expert in their own areas, but this does not make them omniscient."

Peer Review Protocol. After evaluating the survey data, interviews, and observations, the Peer Review Guidance Document (AFCEE, 1996), with minor improvements, is an excellent basis for establishing peer review protocol for the Air Force. The protocol in Figure 17, Recommendation 9, was derived from the Peer Review Guidance Document (AFCEE, 1996) and tailored through observations, interviews, and written feedback. For example, comments made in interviews and on written feedback indicated that the written documents were not utilized before the review. By establishing a protocol that communicates the uses and importance of this written documentation, hopefully the peer review process will be improved.

Future Research

This research developed constructs to measure perceptions of peer review process effectiveness. While process measures are instrumental to continuous improvement, output measures are necessary to verify that the process is actually improved. Current output measures focus on cost savings between the original cost estimate and the cost estimate after peer review. This measure is flawed because it can be easily inflated from

1. Recommend individual peer reviews have a clearly defined scope and specific objectives prior to the review.
2. Recommend peer reviewers be chosen based on the technical parameters of the project(s). For example, if an aquifer is impacted, a hydrologist would be a logical choice for the review team.
3. Recommend each peer review have a designated, trained facilitator.
4. Recommend each peer review have a detailed agenda published prior to the peer review.
5. Recommend project managers utilize an overhead projector and photographs to give a briefing of the overall remediation program and of the specific project(s) being reviewed.
6. Recommend major command peer review coordinators develop a prioritized list of general objectives for the peer review.
7. Recommend the following prioritization for technical peer review focus:
    a) Risk: Is the project necessary based on risk?
    b) Technical: Does the technology appropriate for the media impacted and the risk to human health and the environment?
    c) Cost Effective: Is the technology cost-effective and are costs minimized?
8. Recommend the peer review team serve as a technical advisory board to upper management, not in the pre-approval function.
9. Recommend each major command establish a formal protocol for the peer review process. Propose the peer review process flow as follows:
    a) Project specific scope and objectives along with request for written documentation sent by peer review team leader to project manager.
    b) Project manager sends concise written documentation to peer review team two weeks prior to peer review.
    c) Peer reviewers send requests for clarification to project managers at least three days prior to review.
    d) Peer review team leader publishes agenda at least three days prior to review and distributes to project manager and review team.
    e) Peer review team performs technical review and gives written suggestions to project manager.
    f) Project manager either makes suggested changes to project(s) or makes non-technical rebuttal (e.g., regulatory or public issue) to upper management.
    g) Upper management approves project(s) based on technical peer review approval or non-technical rebuttal from project manager.

Figure 17. Peer Review Process Recommendations

a high original cost estimate and because it does not capture long-term operations and maintenance of installed systems.

There is a requirement to develop an output measure that accurately captures the results of the peer review process. This metric could measure the improvements to the design based on the peer review process objectives: risk, technical, and cost. Constructs need to be developed to capture the degree these objectives are met in the final project. The combination of these constructs would encompass the output metric.

Once the output metric is developed, a longitudinal study could evaluate the effectiveness of various peer reviews. Further, a retrospective study could evaluate projects with no peer review to determine the relative value of peer review. After the output metric has been refined, it could be used to verify the usefulness of the process constructs developed in the current research.

Appendix A

Air Force Peer Review Process Descriptions and Focuses

Three Air Force peer review processes are documented and analyzed: Air Force Center for Environmental Excellence (AFCEE), Air Education and Training Command (AETC), and Air Force Materiel Command (AFMC). The peer review processes are described in the context of how they actually operate versus how policy prescribes they should operate. The three processes and the hierarchy of the top three focuses of each process were described based on documentation, interviews, and observation. The draft processes and focuses were electronically mailed to each perspective program manager for comments and clarifications. In addition, Major Heyse, a member of the AFMC peer review committee, proposed an improvement to the process that may be applicable to other major commands.

Current AFBCA/AFCEE Peer Review Process

Environmental remediation projects of any magnitude are peer reviewed as requested by Air Force major commands using the following process:

1. AFBCA Environmental Coordinator submits Peer Review Information Packages for projects that will be peer reviewed to HQ AFCEE/ERC two weeks prior to peer review.
2. HQ AFCEE/ERC distributes written documentation to peer review team at least one week prior to review.
3. Approximately six technical peer reviewers meet with AFBCA Environmental Coordinator to review projects. HQ AFCEE/ERC uses a dedicated facilitator to moderate the review. A written agenda is recommended. Peer reviewers ask AFBCA environmental coordinator questions about projects that are missing from written documentation.
4. AFBCA Environmental Coordinator makes changes to 1391s and Management Action Plan or Base Closure Plan if required to obtain validation from peer review team.
5. Peer review facilitator completes the Peer Review Validation Form based on consensus from the peer review team.
6. Projects are not funded without Peer Review Validation.

Current AFCEE Peer Review Focus

1. Primary focus: Is the cost minimized?
2. Secondary focus: Is the right technology utilized based on risk?
3. Tertiary focus: Is the rationale adequate for remediation goals?

Current AETC Peer Review Process

Environmental restoration projects requiring peer review are:

- Remedial Designed
- Interim Remedial Actions
- Final Remedial Actions
- Study projects that exceed $500K

The above environmental restoration projects are peer reviewed using the following process:

1. Bases submit Project Summary Reports for projects that will be peer reviewed to HQ AETC/CEVR two weeks prior to peer review.
2. HQ AETC/CEVR distributes written documentation to peer review committee at least one week prior to review.
3. Approximately five technical peer reviewers meet with base program managers to review projects. The Peer Review Committee Chairperson facilitates the review. No written agenda is used, but the Chairperson verbally outlines the process to be used and the primary factors to be discussed -- technical, cost-effectiveness, risk, legal, and political. Base Program Managers present the restoration project(s) justifying the requirement focusing on the primary factors. Peer reviewers ask base program manager questions about projects and provide feedback on aspects of the project lacking justification. The peer review committee also provides feedback on the written

documentation and where improvements are necessary. Peer review committee makes a list of recommendations for changes to the projects.

4. HQ AETC/CEVR submits copy of signed recommendations to base, which is submitted to each Base Civil Engineer and Base Program Manager. The base either concurs with the peer review recommendations and makes the changes or does not concur and raises the issues to the Chief of CEVR for resolution. If Chief of CEVR upholds the peer review committee recommendations, the base must make changes to the project.
5. Base resubmits the Project Summary Report with the necessary changes made.
6. Peer review committee signs modified Project Summary Report (PSR) upon consensus; 100 percent of peer review committee must concur, or PSR is signed by Chief of CEVR. Projects are not funded until PSR is approved and signed by HQ AETC/CEVR.

Current Peer Review Focus

1. Primary focus: Technical
    - Does the technology fit the media impacted?
    - Does the technology fit the risk to human health and the environment?
2. Secondary focus: Cost Effectiveness
    - Is the technology cost-effective or is the cost minimized?
3. Tertiary focus: Legal and Political
    - Is there a legal requirement to conduct remediation or to implement a certain technology?
    - Are there political or public considerations to conduct remediation or to implement a certain technology?

Current AFMC Peer Review Process

All environmental restoration projects that exceed $1M and most projects over $100k are peer reviewed using the following process:

1. Bases submit Project Summary Reports for projects that will be peer reviewed to HQ AFMC/CEVR two weeks ideally (actually may be as little as 2 days) prior to peer review.
2. HQ AFMC/CEVR distributes (sometimes directly sent by base to committee members) written documentation to peer review committee at least one week ideally (sometimes may be as little as 2 days) prior to review.
3. Approximately five technical peer reviewers meet with base program managers to review projects. HQ AFMC/CEVR serves as peer review team leader who facilitates review. No written agenda is used. Base Remedial Program Manager (RPM) presents a case for the project. Peer reviewers ask base RPM questions about projects that are missing from written documentation. Peer review committee recommends the project to HQ AFMC or makes a list of recommendations for changes to projects.
4. HQ AFMC/CEVR submits copy of signed recommendations to base. Base either concurs with recommendations and makes changes or does not concur and raises objections to the peer review team leader, who may discuss them further with the committee or take the issues to the Chief of CEVR for resolution. If Chief of CEVR upholds peer review committee recommendations, base must make changes to project before funds are released.

5. Base resubmits the Project Summary Report with the necessary changes made.
6. Peer review committee signs modified Project Summary Report (PSR) upon consensus; 100 percent of peer review committee must concur, or PSR is signed by Chief of CEVR. Projects are not funded until PSR is approved and signed by HQ AFMC/CEVR.

Current AFMC Peer Review Focus

1. Primary focus: Is project needed based on risk?
2. Secondary focus: Is the right technology utilized?
3. Tertiary focus: Is the cost minimized?

Proposed Modification to AFMC Process (Maj Heyse)

1. Base submits Project Summary Report two weeks prior to peer review to HQ AFMC/CEVR and the peer review committee.
2. Approximately four to five peer reviewers meet with base program managers to review projects. HQ AFMC/CEVR facilitates review. No written agenda is used. Peer reviewers ask base program manager questions about projects that are missing from written documentation.
3. Peer review team signs list of recommendations and HQ AFMC/CEVR submits list of recommendations to base; consensus from peer review team members is sought but not required.
4. Base then endorses the peer review committee recommendations and concurs/nonconcurs with each recommendation. Base provides a rebuttal to issues which they nonconcur.
5. Base submits endorsed peer review committee recommendations to HQ AFMC/CEVR. HQ AFMC/CEVR attaches recommendations with base endorsement to original Project Summary Report. Project Summary Report is not modified.

Appendix B

Air Force Base Conversion Agency Peer Review Observations

29 Feb 96 Observation of Peer Review of a closed Air Force Base in California

Background. There were 20 people present at the peer review. AFCEE hired a contractor, Waste Policy Institute, to facilitate the peer review. There were eight people on the peer review team including the facilitator. Three people represented the projects at Castle. Observers included one state regulator and one Environmental Protection Agency regulator.

Observations. The peer review team did not have time to read the project documentation prior to the review. The project documentation included approximately 13 packages with a combined total thickness of approximately five inches. The program manager appeared to be irritated that the peer review team had not read the documentation and was delving into details. One comment was, "Let's go ahead or we'll be here seven days." One person on the peer review team said, "I am very confused about what we are supposed to be doing here. What are we validating?" The program manager stated, "We are prematurely reviewing this package." The peer review team responded by asking if there were any projects that were "ripe for funding." The facilitator emphasized that the goal was to validate technology and costs. The program manager asked if this was the first time the peer review team had reviewed the written documents. Several people on the peer review team responded, "Yes." The program manager stated that they will not have the needed data until August 1996 and that it is “iffy” trying to evaluate costs and technology before data is compiled and decisions are made. A peer review team member stated that the objective of the peer review is for program managers to convince the team that the projects are valid. The program manager stated that last year the peer review team had not read the RI/FS either. A peer review team member stated that, “We are trying to see the forest through the trees.” The facilitator stated, “Lets focus on choosing the technology; the costs will fall out.” The program manager stated that “This is not a salesman issue;” this should only take 2-3 hours since there is not enough data. The Project Summary Sheets were approved/disapproved by peer review team vote.

29 Feb 96 Observation of Peer Review of a Base Realignment and Closure listed Air Force Base in Texas

Background. There were 26 people present at the peer review. AFCEE hired a contractor, Waste Policy Institute, to facilitate the peer review. There were eight people on the peer review team including the facilitator. Five people represented the base.

Observations. Base presenters utilized overhead maps with colored legends which helped observers visualize the sites.

Analysis

The following analysis is based on observations of the Air Force Base Conversion Agency peer reviews:

1. The peer review team did not have sufficient time to read documentation prior to the review.
2. The project documentation could be more clearly organized and more brief.
3. The program manager did not provide a good overview that "sold" the projects.
4. An overhead projector was available and was helpful when used.
5. Name tags and placards on the table to identify peer review participants would be helpful.
6. Peer reviewer and facilitator people skills were instrumental in making the process effective despite of an adversarial climate.
7. Several peer review team members made comments that a cost analyst was needed on the team.
8. The process would have been more effective if there was a clearly established focus and agenda established prior to the review.

Appendix C

Peer Review Interviews

Figure 18 shows the interview protocol used throughout the research. Interviews were in person, when possible, and by telephone otherwise.

1. Give your whole attention to the person interviewed.
2. Listen--don't talk.
3. Never argue--never give advice.
4. Listen to the interviewee.
5. As you listen, plot out tentatively and for subsequent correction, the pattern that is being set before you. To test this, occasionally summarize what has been said, and present it for comment. Always do this with caution; that is, abbreviate and clarify but do not add or 'twist.'

Figure 18. Interview Protocol (Adapted from Doraiswami, 1963:46)

9 Oct 95--In-person interview with Maj Heyse (AFMC Peer Reviewer)

- Discussed general characteristics of AFMC peer review program.
- Suggested review what other major commands are doing for peer review.

17 Oct 95--In-person interview with Mr. Wayne Ratliff (AFMC Peer Review Program Manager)

- Discussed AFMC Peer Review Program--Mr. Ratliff is POC.
- The Headquarters of the United States Air Force Environmental Restoration Division, noticed problem with IRP projects and initiated the peer review program.

- Recommendations:

    1. Look at peer review process.
    2. Is the peer review process a paperwork exercise?
    3. Do bases follow peer review advice?
    4. Survey bases under peer review.
    5. Survey major commands.

2 Nov 95--Telephone Interview with Mr. Bob Furlong (AF/CEVR)

- Peer review started in 1992 by AF/CEVR, Col Owendorf.
- Looking for methods to improve cross-feed of peer review information between major commands.
- Annual Air Staff guidance delegates responsibility for peer review programs to major commands.
- AFCEE uses in-house and contractor to peer review base closure.

7 Dec 95--In-person interview with Mr. Wes Lanen (Elmendorf RPM)

- Peer review can help filter regulatory and political biases for the Remedial Project Manager (RPM).

11 Dec 95--Telephone interview with Mr. Larry Jackson (Waste Policy Institute. Completed Masters Thesis on Peer Review. Studied PACAF, AETC, AFMC, and ACC.

- ACC completes peer review by contract.

- AETC uses college professors to complete peer review.
- AFMC completes peer review in-house.
- PACAF uses EPA, in-house, and other major command staff to conduct peer reviews.
- Possible research question: Are major commands spending more time and money than the peer review is worth?
- Mr. Jackson compiled characteristics of the four major command's peer review programs then asked them to concur.
- Telephone interviews more successful than written.
- Three out of four major commands had written peer review protocols.
- Possible metric: How many projects reviewed? How many projects revised?
- From thesis, he recommends major commands develop standardized guidelines.
- Need distillation of best practices.
- Avoid monetary metric.

18 Jan 96--In-person interview with Maj Nixon (AFMC peer review member)

- Compare peer reviewed projects to non-peer reviewed projects.
- How is it best to do peer review?
- Is it beneficial to do peer review? (Without question "everyone" says yes.)
- What is the function of peer review?
- What are the boundaries of peer review?
- What variables should be considered in evaluating the peer review process?

- How many projects per day is the right balance between efficiency and effectiveness?
- Financial vs. scientific vs. political
- Experience mix of team members is important.
- Survey members of peer review committee and PMs who have been exposed to peer reviews.
- How is peer review linked to decision making process?

18 Jan 96--Telephone interview with Mr. Bill Lopp (AETC Peer Review Program Manager)

- Cookbook approach guidance to hand-off to project managers for peer reviews would be useful.
- Define peer review.
- Establish guidelines for how a peer review should be conducted.
- Establish a checklist of important features for peer reviews.
- * What is the measurement of success of project peer reviews?
- IRP Projects exceeding $1M are peer reviewed.
- What do the Project Managers need to do?
- Ideas on a measurement of success for peer reviews:
  - Recommendation of peer review to modify proposed plan or feasibility study.
  - Did peer review save money, resources, time?
  - Did peer review evaluate presumptive remedy (Insitu)?

- We need a peer review data base.
- Need a guidance document, "This is what a good peer review looks like."
- Peer review protocol should be modified for different media: groundwater, soils, etc.

1 Feb 96--In-person interview with Maj Dan Welch (AFCEE/ERC)

- This week sent out questionnaire to major commands getting peer review guidance input. AFCEE is taking the lead to establish guidelines for the Air Force. AFCEE is drafting peer review guidance to be implemented by Air Staff.
- AFCEE/ERT (Technology Transfer) currently collects data on IRP projects and sends new technology information into the field.
- AFCEE peer review includes one member from Technology Transfer on each peer review team.
- AFCEE has not established metric but recommends: Number of projects peer reviewed versus estimated dollars saved. [Value engineering approach?]

1 Feb 96--In-person interview with Mr. Jerry Hansen (AFCEE/ERT Technology Transfer)

- There are approximately five people assigned to Technology Transfer.
- He has been on two peer reviews and thinks they are a good process.
- Hiring a contractor to conduct peer reviews is probably gaming the system.

1 Feb 96--In-person interview with Mr. Dave Cook (PACAF/CEVR)

- Timing is critical to a successful peer review process.
- They tried video teleconference peer reviews and found that they weren't so good.
- Recommends conducting approximately five peer reviews in a row (one per day) to save travel money.

1 Feb 96--In-person interview with Lt Col Ross Miller

- Peer review was initiated when landcaps were considered a presumptive remedy and were designed underwater--which is not effective--in several BRAC locations.
- Mr. Gary Vest made decision to broaden peer review to IRP and charged Col Owendorf with implementing on active bases.

1 Feb 96--In-person interview with Col Selstrom (AF/CEVR)

- Will be forming Tiger Team in the Spring to look at proposed guidance to implement peer review at the major commands.

1 Feb 96--Discussion from "floor" during presentation by Maj Dan Welch (AFCEE/ERC) on the peer review process. Mr. Bob Moore (ACC/CEVER) and Mr. Bob Furlong (AF/CEVR)

- Mr. Moore: ACC's peer review process involves contractor reviewing IRP project documentation without base present. Base representatives are only brought in on non-concurrances.
- Mr. Furlong: We need a "gut check" [peer review] because the Air Force is spending a lot on RAs [Remedial Actions]. If AFCEE designs project, we do not need peer review.
- Mr. Moore: I still think AFCEE projects need peer review.

1 Feb 96--In-person interview with Mr. Bob Furlong (AF/CEVR0

- Metrics for peer review are currently not available. Installing the best technology is more important than saving money.

1 Feb 96--In-person interview with Mr. Mike Bataglia (ACC/CEVR)

- Peer reviews currently cost approximately $7k-$10k per review.
- Does not like the idea of having one AF office oversee another AF office within the same organization.

2 Feb 96--In-person interview with Mr. Bill Lopp (AETC/CEVR)

- AETC peer reviews all projects with estimates > $500k.

2 Feb 96--In-person interview with Ms. Diane Butlak (SPC/CEVR)

- Peer review should be early in the IRP process.
- Peer review should take into consideration political issues.

29 Feb 96--In-person interview with Mr. Cesar Silva (AFCEE/ERB)

The following are important issues for a successful peer review:

- Documentation should be provided to peer reviewers with adequate time for a thorough review prior to the meeting.
- The peer review team should have the "right" mix of specialists.
- Good representation from program manager; they should be able to justify what they are doing.
- McClellan Air Force Base program managers used color overheads, maps, and a very detailed briefing. They had regulator "buy in" on programs presented.
- Regulators were not included at Castle, no "buy in." The preliminary documents were not provided to regulators prior to the Castle peer review.
- The peer review should focus on three areas: costing, risk assessment, and technology.

29 Feb 96--In-person interview with Dr. Bill Sweet (AFCEE/ERC)

- Four specialists are needed on the peer review team: hydrologist, toxicologist, chemist, and an environmental engineer/cost analyst.

- We need to develop a way to inform bases what is needed for a peer review.
- Program managers should give a 10-20 minute presentation including a site conceptual model, demographics, land use patterns, regional topology, and surface and ground water characteristics.
- Program managers should briefly describe each project including: important constituents (politically, legally, or risk based), intended land use, risk assessment, proposed remedy, rational, and cost evaluation.
- The program manager should convince me that the project is good and rational.

29 Feb 96--In-person interview with Ms. Barbara Smith-Townsen (AFCEE/ERC)

- McCellan AFB peer review was excellent. They provided a good package [documentation] with excellent organization; they were prepared; they had data organized for easy access during their briefing.
- They grouped sites into programs versus projects.

29 Feb 96--In-person interview with Ms. Judy Strickland (AFCEE/ERC)

- Peer review committee should be trimmed to one type of specialist on each team. Duplicate expertise such as two chemists and three geologists is not needed.
- We need cost experts on peer review teams.
- Facilitator should keep peer review focused.
- There should be an agenda.

- There should be better communication of what is expected from reviewers to program managers.
- There was not enough information on risk.
- The metric for a good peer review should be how many technologies changed and how much money was saved.

29 Feb 96 and 1 Mar 96--In-person interview with Capt Mark Smallwood (AFBCA Peer Review Program Manager)

- A facilitator is important to the peer review process.
- Should all projects be peer reviewed?
- We need a framework for systematic review of projects to optimize usage of time.
- Having a schedule is important.
- Need a strong facilitator to keep the peer review team on track and focused.
- A lot of variables are not technical: regulatory, political, community, reuse, etc. Any of these can drive requirements.
- Don't want to look at details such as depth of wells, screening, etc.
- Do want to review assumptions and the broad perspective.
- Types of peer review questions: "How often do you sample? When did you last review the number of wells?
- A peer review is a "stand back look," and a "sanity check."
- Rely on existing data during peer review. Do not open up RAs etc.

- Out of cycle peer reviews versus annual. Annual peer reviews have a focused effort, technical experts, EPA labs, etc.
- Measure of merit (metric): Reduction in requirements being programmed.

2 May 96--Telephone interview with Mr. Dave Stokes (AETC/CEVR)

- Peer review should focus on root causes: (1) What does the process do to detect a poor submission to peer review, (2) What does the process do to fix that problem?
- Fix front-end instead of back-end.

7 Aug 96--Discussed AFMC Peer Review Process with Maj Heyse and Maj Nixon

See Appendix A for AFMC Peer Review Process.

7 Aug 96--Discussed AETC Peer Review Process with Mr. Richard Trevino

See Appendix A for AETC Peer Review Process.

Appendix D

Air Force Base Conversion Agency Peer Review Feedback Forms

Background

The Air Force Base Conversion Agency peer reviews, administered by the Air Force Center for Environmental Excellence, were held from 27 February to 6 March 1996. The projects under review were from the following installations: McClellan, Lowry, Myrtle Beach, Rickenbacker, Loring, Kelly, K.I. Sawyer, Castle, Wurtsmith, Mather, Plattsburg, Reese, March, Chanute, and Norton. Feedback forms were available and completed throughout the duration of the peer reviews. Twenty two Feedback forms were completed. The notes and comments are reproduced as hand written with the exception of spelling errors. Information in brackets "[ ]" is added for clarity.

The following title and instructions were printed on each form:

> "PEER REVIEW NOTES and COMMENTS
> Please use this page to make comments and lessons learned ideas
> which will assist us next year to improve the Peer Review."

Feedback data is numbered one through twenty-two below.

1.

- Much of the dollar requirements are contingent upon regulator issue resolution.
- Many of these requirements become firm and an addition look at these projects may be necessary.

2.

General Comments

If the projects can not be canceled, as it stops the funding request[,] and the incomplete technical information are provided, the only alternative is to recommend by peer review committee "Validate alternative technology and cost," for most of the projects.

3.
Devote 1st 2-3 hours to detailed presentation of Base program status then project review rather than going straight into first project, second project, etc....

4.
It would help tremendously to have peer review pkgs. QC [quality control] prior to the weeks of actual review. Some info were missing. It took time to dig out those info such as gw [ground water] contour map form actual report.

5.
SUGGESTIONS

Do one per base Myrtle Beach
(1) provide project life cycle cost (PV) rather than only cost for next FY
(2) provide technical + cost feasibility study of alternative remedies whenever possible

6.
Imperative that presenter introduce the base: base-wide geology, hydrology, (surface) hydrology, historical activities at the base, and proposed uses. Same thing for site-specific information: succinctly but specifically cover geology, hydrology, site activities, proposed uses, chemical concentrations (if available) and risk numbers (analytes and associated non-carcinogenic and carcinogenic risk).

7.
(1) Include Brac Peer Review comments in 1391.
(2) Recommend use of re-analyzed analyte data, instead of J-flagged estimated data for decision making and cost calculations.

8.
Suggestions to Improve Process
(1) Provide relevant information far ahead of time.
(2) Provide a short overview of project by overheads.
(3) Provide a 5/10 minute break after Presentation through overheads to digest information (if no:1 is not provided)

9.
Cal RWQLB [California State Regulator]

- All participants read to be prepared. The peer reviewer should have read the documents. The base should have a prepared presentation.
- Mostly I feel this process was unproductive and unorganized. The advice given by peer reviewer has little useful value without more preparation so they can understand general concepts of the project. I feel this is a waste of government resources.

10.

Peer Review Comment

(1) Some individuals do not have background or experience with remediation technologies. Could there be a pre-peer review seminar available for those that could attend or review the various technologies. We are asked to evaluate things we don't have any idea about.

(2) Pls have info available for various bases. If at all possible, maybe a video presentation of a summary for the base, etc.

(3) Facilitator was outstanding - he should train the others in the methods of facilitatorship.

11.

(1) Process developed into a good data exchange

(2) Improvement

Base Present 15 Min prepared presentation

- History of IRP
- Where we are now
- Where we agree w BLT
  " " disagree w BLT

(3) Review Projects

As this year present project descriptions early but distribute ahead of time to team.

Base will have program book in order and correlated with descriptions.

Keep same team

12.

- The primary drivers for why a project is being undertaken are implicit in the presentations: however, it would be helpful if right up front the actual driver for this project was stated + then the background were given. Also an identification of any secondary drivers {ARARs?} [Applicable or Relevant and Appropriate Requirements] that "force" an issue being upfront would also be good.
- This presentation presented "the facts" from which we as a group needed to decide if it should proceed without for knowledge of the above conditions.

13.

- Part of this review is the evaluation of cost. However, several times several of the panel experts stated that they don't know anything about cost or how to estimate the cost. If this is to continue to be an important part of this review we need to include experts (or at least some informed people) in this area.
- We need to get these review packages to the reviewers before the meeting. A quality review of each of the projects can't be done in 10-15 minutes.

14.

Peer Review Suggestions

(1) Provide specific instructions to BCA-OL [BRAC project manager] that a presentation will be required.
(2) Have a peer cost analyst available during session.
(3) [AFCEE/] ERC needs to proceed in a much more organized manner (they did not provide a good picture of AF in this mtg in front of regulator agencies)
(4) Provide mechanism for including regulators. They feel it is a waste of their time.
(5) Chris did a good job as facilitator after the initial chaotic start.
(6) Not much input/output from some ERC staff.
(7) Why isn't ERT involved when RD/RA issues are being discussed[?]
(8) Provide clear agenda for all attendees.

15.

Proposed Improvements Feb 96 SA Peer Rev.
(1) Peer review team review submitted pkgs. prior to mtg.
(2) Peer review team review base RI/FS reports prior to mtg.
(3) Base provide overall general presentation prior to site specifics.
(4) Base receive call for peer review packages ≈ 3 months prior to mtg.
(5) Peer rev team receive & review base packages 1 mo. prior to mtg.

Items that went well
(1) Professional demeanor maintained
(2) Pace and duration

16.

(1) Conduct a pre-review or dry run at the installation with AFCEE and AFBCA personnel.
(2) Standardize and simplify project summary sheets.
(3) Regulators did not contribute anything. Do they really need to be there?
(4) Facilitator performed well and greatly enhanced the meeting productivity (Thanks to Chris Miller of WPI)
(5) The burden of educating the PEER group on each project should be the responsibility of the operating location (OL). Likewise, it is the responsibility of HQ AFBCA to provide adequate direction and guidance to the OL on what they need to provide the group.
(6) Perhaps, instead of a dry run at the base, the base could come to AFCEE two weeks prior to present materials to PEER group, then return in two weeks for the review.

17.

Castle

- Review materials must be given to reviewers at least one week before the date of review.
- Simplify review materials with everything in one package for each program for which money is required.
- Dry run the presentation to ensure that all materials are correct & presenter is prepared.

- Make view graphs.

18.
- Rickenbacker ANG - The meeting room was too small for the large group.
- Three chemists, two from [AFCEE/]ERC and one from AL were on the review team--obvious overkill. Lack of communication between BA&H and ERC concerning distribution of site background data.
- General--lack of adequate restroom facilities!!
- Recommend that ratio of diet soda:sugar soda be increased next year.
- Recommend that a large-scale map of each base showing IRP sites and other project sites be provided in each room for each Base.
- Recommend that no peer review member be asked to be the recorder. -This position should be filled by a rep from BA&H, WPI[contractor] or AFBCA.

19.
- The peer review process in its present form is not that useful or productive. This process would be better served by technical audit teams that visit each site, review documents before asking questions and have access to all site data.

20.
- Use quarterly BEC conference to give requirements for the peer review.
- Develop model peer review package to help bases assemble their materials. Key examples to BRC's needs and objectives. Provide to reviewers also, well in advance. There is still much misunderstanding on the objectives of the Review and needs of the review team.
- Schedule further in advance.
- Provide feedback to BECs/other BRAC presenters-reward good presentations, counsel poor presenters.

21.
- Costs/technologies cannot always be fully evaluated/validated due to uncertainties in "required" activities in early stages of process (i.e., pre-decision information)
- 1391's narrative portions do not offer enough information for peer reviewers to fully evaluate projects (e.g., list of specific COC's, better descriptions of sites-resources, range of concentrations measured, reuse considerations). Additionally, conceptual site models (were not presented), maps (depicting not only site locations but installation characteristics and/or major environmental features) should be requested from installation to be presented in conjunction with 1391's. AFCEE should coordinate w/ installations prior to peer review to clarify document needs.
- Bioremediation and/or natural attenuation although a site-specific valid technological consideration does not always meet the realistic demands for property transfer under BRAC goals.

- Technical reviewers do not always understand the environmental restoration process (RI/FS, PA/SI, RA, RD, etc.) Side discussions concerning these issues distract form the technological considerations. The reviewers briefings or briefing materials should contain this information.
- The BCI's should make clear what are clearly regulatory requirements so as not to mislead the technical reviewers into debating whether a project should be required, e.g., whether a project should be validated based on a technical reviewers personal opinion of unacceptable risk.

22.

- AF needs to push for putting restrictions of future use in deeds to property or not transfer.
- AF needs to push for putting restriction of future use in deeds to property or not transfer to another.
- Landfills(esp. Municipal) should be treated as Bioreactors and monitored and hydraulically controlled but not capped.
- Lowry's landfill costs then should go from $22M to less than $3 million.
- Let the bases see each other's 1391 and handouts. Many have same type sites and could benefit from each others.
- Lowry needed to give us more documentation on contamination levels for us to evaluate technical decisions.
- Next year have an Environmental Economist in each room to validate costs.

Analysis.

Many issues were raised through the peer review feedback forms. The most common comment addressed the importance of the presentation of the projects under review; 12 out of the 22 feedback forms addressed the presentation. Next in importance is clearly written advanced documents describing the projects; 8 out of 22 feedback forms addressed advanced material. Four respondents noted that there was a problem with the mix of specialists on the peer review team. Three thought there should have been a cost analyst and one noted that there were three chemists on their peer review team, in their words, "overkill." The focus of the peer review was mentioned; both regulatory and funding issues were addressed. Lastly, the importance of an agenda, a good facilitator, and peer reviewer skills were addressed.

Appendix E

Questionnaire

The questionnaire that follows was distributed to people who participated in the Air Force Center for Environmental Excellence (AFCEE) administered peer reviews, the Air Education and Training Command (AETC) peer reviews, and the Air Force Materiel Command (AFMC) peer reviews. The cover letter on page E-2 was sent to people participating in the AFCEE peer review. The cover letter on page E-8 was sent to people participating in the AETC peer review. For AFMC, the questionnaires were administered in-person.

Letterhead

MEMORANDUM FOR PEER REVIEW PARTICIPANTS

FROM: HQ AFCEE/ERC

SUBJECT: Peer Review Survey

1. Background: DoD and USAF Air Staff initiatives direct that peer reviews be conducted for environmental remediation projects or environmental studies prior to consideration in the annual funding program. We are appointed as the Air Force OPR to develop and maintain the Peer Review Program..

2. Problem: The Peer Review Program has been instituted across the Air Force at all Major Commands (MAJCOMs) and the Air Force Base Closure Agency (AFBCA). The Peer Review Program varies widely between different MAJCOMs and the AFBCA.

3. Purpose of Survey: The purpose of this survey is to refine the Peer Review Program. There are no right or wrong answers. This survey is being conducted for us by the Air Force Institute of Technology.

4. Confidentiality: Participation in this survey is voluntary, results will be reported aggregately.

5. Questions: Please address questions to Capt Paul Schantz, email: pschantz@afit.af.mil or voicemail DSN: 785-3636, ext 6395. I solicit your prompt cooperation in this project and thank you for your time.

[Left Blank], Lt Col, USAF
Chief, Consultant Flight

2 Attachments:
1. Instructions
2. Survey

The attached survey is designed to be completed on a specific peer review.

PLEASE ANSWER SURVEY QUESTIONS BASED ON

________________________ AFB PEER REVIEW

CONDUCTED ON ______________________ 96.

Return the completed survey by 30 Jun 96 or complete the survey in black ink and fax to: (513) 476-4699 or DSN: 986-4699.

CIRCLE ONLY ONE NUMBER PER QUESTION, the response that applies best to you.

GROUP PROCESS	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
1. The Base Program Manager presented adequately detailed funding estimates of project(s) under review.	1	2	3	4	5
2. The Base Program Manager's presentation of *proposed* remedial actions was clear.	1	2	3	4	5
3. The Base Program Manager's presentation of *alternate* remedial actions was clear.	1	2	3	4	5
4. The Base Program Manager's presentation of rationale to meet regulatory requirements was understandable.	1	2	3	4	5
5. The Base Program Manager's presentation of risk-based approach was clear.	1	2	3	4	5
6. The Peer Review Team completely reviewed the technical merits of all *proposed* remedial actions.	1	2	3	4	5
7. The Peer Review Team completely reviewed the technical merits of all *alternate* remedial actions.	1	2	3	4	5
8. The Peer Review Team completely reviewed the Base Program Manager's rationale to meet regulatory requirements.	1	2	3	4	5
9. The Peer Review Team completely reviewed the Base Program Manager's risk assessment results.	1	2	3	4	5
10. The Peer Review Team adequately reviewed funding estimates.	1	2	3	4	5
11. This overall Peer Review Process was effective.	1	2	3	4	5
12. This Peer Review was worth the time and expense required to conduct it.	1	2	3	4	5
FOCUS					
13. The desired end-result of the Peer Review was stated with precision.	1	2	3	4	5
14. The scope of the Peer Review was well defined.	1	2	3	4	5
15. The Peer Review was focused on validating funding estimates.	1	2	3	4	5
16. The Peer Review was focused on cost savings.	1	2	3	4	5
17. The Peer Review was focused on evaluating the Base Program Manager's rationale to meet regulatory requirements.	1	2	3	4	5
18. The Peer Review was focused on validating *assumptions* of the projects.	1	2	3	4	5
19. The Peer Review was focused on validating *design* of the projects.	1	2	3	4	5
20. The Peer Review included a review of all non-standard *assumptions*.	1	2	3	4	5

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
FOCUS (continued)					
21. The Peer Review included a review of all non-standard *designs*.	1	2	3	4	5
22. The Peer Review focused on scientific and technical aspects of the projects.	1	2	3	4	5
23. The Peer Review was focused on evaluating alternate technologies.	1	2	3	4	5
24. The Peer Review was focused on evaluating risk assessment results.	1	2	3	4	5
AGENDA & FACILITATOR					
If there was not a planned agenda for the Peer Review, please skip to question 28.					
25. The Peer Review closely followed the agenda.	1	2	3	4	5
26. The Peer Review agenda was easy to follow.	1	2	3	4	5
27. The Base Program Manager had adequate time to introduce the projects under review.	1	2	3	4	5
28. If there was not a facilitator for the Peer Review, please skip to question 29. The facilitator kept the Peer Review focused.	1	2	3	4	5
WRITTEN PREPARATION					
29. The Peer Review Team had adequate time to review documentation on projects under review provided by the Base Program Manager.	1	2	3	4	5
30. The Peer Review Team read documentation provided by the Base Program Manager before the Peer Review.	1	2	3	4	5
31. The written material provided by the Base Program Manager was organized for quick reference.	1	2	3	4	5
32. The Base Program Manager provided written material at the right level of detail for efficient review.	1	2	3	4	5
ORAL PRESENTATION					
33. The Base Program Manager's briefing was at the right level of detail for efficient review.	1	2	3	4	5
34. The Base Program Manager had detailed project data readily available to answer Peer Review Team questions.	1	2	3	4	5
35. The Base Program Manager clearly described the project location(s).	1	2	3	4	5
36. The Base Program Manager briefed a clear description of the overall remediation program.	1	2	3	4	5
37. The Base Program Manager briefed a clear description of individual projects under review.	1	2	3	4	5
38. The Base Program Manager "sold" the project to the Peer Review Team.	1	2	3	4	5

PRESENTATION (continued)

	Far Too High	Somewhat High	OK	Somewhat Low	Far Too Low
39. The level of detail of project documents provided by the Base Program Manager was:	1	2	3	4	5
40. The level of detail of briefings presented by the Base Program Manager was:	1	2	3	4	5
PEER REVIEW TEAM CHARACTERISTICS					
41. The number of specialists on the Peer Review Team was:	1	2	3	4	5
42. The experience of specialists on the Peer Review Team was	1	2	3	4	5

PEER REVIEW TEAM CHARACTERISTICS (continued)

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
43. The Peer Review Team was objective.	1	2	3	4	5
44. The Peer Review Team was constructive.	1	2	3	4	5
45. The Peer Review Team adequately considered community concerns.	1	2	3	4	5
46. The mix of specialists on the Peer Review Team was right.	1	2	3	4	5
REVIEWER CHARACTERISTICS					
47. The Peer Reviewers were biased.	1	2	3	4	5
48. The Peer Reviewers were independent of the projects being reviewed.	1	2	3	4	5
49. The Peer Reviewers had adequate knowledge to perform a thorough Peer Review.	1	2	3	4	5
50. The Peer Reviewers were good listeners.	1	2	3	4	5
51. The Peer Reviewers were good interviewers.	1	2	3	4	5
52. The Peer Reviewers were good communicators.	1	2	3	4	5

JOB INFORMATION

53. Which Major Command or Agency are you employed or contracted by?

1. HQ USAF (Air Staff)
2. ACC (Air Combat Command)
3. AETC (Air Education & Training Command)
4. AFCEE (Air Force Center for Environmental Excellence)
5. AFMC (Air Force Materiel Command)
6. AFSPC (Air Force Space Command)
7. AMC (Air Mobility Command)
8. PACAF (Pacific Air Forces)
9. AFBCA (Air Force Base Closure Agency)
10. OTHER: (Please write in) ____

54. What *role* did you primarily play in this Peer Review?

1. MAJOR COMMAND OR AIR STAFF OVERSIGHT
2. BASE PROGRAM MANAGER OR BASE REALIGNMENT AND CLOSURE ENVIRONMENTAL COORDINATOR
3. FACILITATOR
4. PEER REVIEWER--TECHNOLOGY TRANSFER
5. PEER REVIEWER--ENGINEER OR SCIENTIST
6. OTHER: (Please write in) ____

OPEN ENDED QUESTIONS:

55. The following type of specialist should have been on the Peer Review Team: ______________________.

56. The Peer Review Process would be improved by...____

____

____

____

57. This survey would be improved by...____

____

____

____

58. This survey should have asked...____

____

____

____

MEMORANDUM FOR AIG 10904

FROM: HQ AETC/CEVR
266 F Street West
Randolph AFB TX 78150-4321

SUBJECT: Peer Review Survey (SUSPENSE: 18 Jul 96)

1. As we discussed during our Video Teleconference (VTC) on 27 Jun, attached is the Peer Review Survey. Please complete the attached survey and return it to us by 18 Jul, COB.

2. The survey should be completed by bases that participated in the Peer Review process. The intent of the survey is to provide constructive feedback on the HQ AETC/CEVR Peer Review process, thereby improving the process for the future.

3. If there are any questions, please do not hesitate to contact Richard Trevino, DSN 487-3302.

Chief, Environmental Restoration Branch
Environmental Management Division
Directorate of The Civil Engineer

Attachment:
Peer Review Survey

Appendix F

Survey Development Procedure

Survey development began with a definition of the purpose. Establishing the purpose was the first step in the fifteen step methodology used to develop the survey (Figure 18).

1. Define the purpose. Be specific!
2. Review existing data. was a survey needed?
3. Read applicable regulations.
4. Define the hypothesis.
5. Define the population.
6. Develop the survey (& sampling) plan.
7. Develop cover letter, instructions, & Privacy Act Statement.
8. Develop survey questions.
9. Pretest instrument.
10. Edit and revise questionnaire.
11. Obtain approvals as required.
12. Survey (gather data).
13. Quality control/data reduction.
14. Analysis and interpretation of results.
15. Prepare report for customer(s).

Figure 19. Steps in Survey Instrument Design (Adapted from DoAF, 1993:47).

In the second step, the existing data was reviewed from the literature review, interviews, and observations to determine if a survey was needed. Also at this point, the purpose and data deficiencies were evaluated to ensure a survey was the best tool for meeting the research objectives and questions. Third, applicable federal regulations were reviewed to determine what procedural steps must be taken to administer the survey. Air Force Instruction 36-2601--*Air Force Personnel Survey Program*--and Air Force Instruction 37-124--*The Information Collections and Reports Management Program; Controlling*

Internal, Public, and Interagency Air Force Information Collections--were reviewed to determine specific procedural requirements and constraints. Since the approach used in this research was a non-experimental design, there was not a hypothesis rather the relationship between the six constructs and peer review effectiveness was studied (Hatcher, 1994:9).

In the fifth step, the population of the study was defined. Next, the survey plan was developed which includes the data collection plan, the data reduction plan, and the analysis plan (DoAF, 1993a:11). The survey cover letter and instructions were developed next and an attempt was made to get the highest level of sponsorship possible. A twelve step process (DoAF, 1993a:31-34) was used to develop the questionnaire.

The Air University Sampling and Surveying Handbook gives a list of twelve guidelines which were used to develop the survey questions (Figure 19).

1. Keep the language simple.
2. Keep the questions short.
3. Keep the number of questions short.
4. Limit each question to one idea or concept.
5. Do not ask leading questions.
6. Use subjective terms such as good, fair, and bad sparingly, if at all.
7. Allow for all possible answers.
8. Avoid emotional or morally charged questions.
9. Understand the should-would question.
10. Formulate your questions and answers to obtain exact information and to minimize confusion.
11. Include a few questions that can serve as checks on the accuracy and consistency of the answers as a whole.
12. Organize the patterns of questions.

Figure 20. Guidelines for Survey Question Development (Adapted from DoAF, 1993a:31-34).

These guidelines along with those presented by Babbie (1990:127-132) were used throughout survey development to improve the quality of each question. Survey items were developed from characteristics of constructs into statements. These statements were reviewed to ensure the language was simple and the statements were as short as possible. Further, each statement was reviewed to ensure they were not testing multiple ideas. Leading questions bias survey results and were removed from the survey. The language used on the survey was as objective as possible to reduce subjectivity.

The survey items were further refined to ensure the responses allow for all possible answers. Each survey item was worded to address specific characteristics of the peer review process. Survey questions were organized by the constructs which the items were testing. To check the accuracy of the responses, one reverse coded question was included. Once the survey was constructed, the next step was to pretest the survey to ensure the questions were clear and logical. Following the pretest, the survey was edited to correct noted problems. The corrected survey was then forwarded for approval by necessary Air Force offices, if required.

The Air University Sampling and Surveying Handbook (DoAF, 1993a) recommends six techniques to increase survey response rate (Figure 20).

1. Use follow-up letters.
2. Use high-level sponsorship.
3. Make your questionnaire attractive, simple to fill out, and easy to read.
4. Keep the questionnaire as short as possible.
5. Use your cover letter to motivate the person to return your questionnaire.
6. Use inducements to encourage a reply.

Figure 21. Six Techniques to Increase Survey Response Rate (Adapted from DoAF, 1993a:37-38

Follow-up letters or phone calls are also recommended to increase response rate; however, these methodology were not used due to the necessary time lapse in administering the survey.

The twelfth step was to administer the survey to the sample chosen from the population. Upon receipt of the results, the data was reviewed for quality control and data reduction. Once the data was 'cleaned,' it was analyzed and condensed into results. The final step was to prepare the results for the customer; in this research, the final step was to relate the results back to the research objectives and questions and draw conclusions.

Appendix G

Survey Administration Approval Requirements

Read Applicable Regulations

Air Force Instructions (AFIs) 36-2601 and 37-124 are the two regulations that pertain to administering an Air Force survey. AFI 36-2601 provides guidance for getting approval and conducting surveys within the Air Force. AFI 37-124 is a more general regulation that establishes guidelines for managing and controlling information collection and reporting within the Air Force, to the public, and interagency. AFI 37-124 specifically requires approval for certain information requests by the Office of Management and Budget (OMB).

AFI 36-2601 was reviewed thoroughly and applied to the research sample. The research sample came from people working on the Air Force Base Conversion Agency project peer review. These people fell under the following categories: people assigned to Brooks AFB, people assigned to the Air Force Base Conversion Agency, Environmental Protection Agency Laboratory employees, state Environmental Protection Agency Employees, and contractors. Each category was reviewed to determine whether approval is required, and what steps were required.

For personnel assigned to Brooks AFB, exception 1.6 of AFI 36-2601 applies, "Single-base surveys initiated by the installation or unit commander only on issues under his or her control." Since the Air Force Center For Environmental Excellence (AFCEE) is responsible for executing the Air Force Base Conversion Agency project peer review, AFCEE approval, through survey cover letter signature, meets the requirements of AFI 36-2601 through exception 1.6.

For personnel assigned to the Base Realignment and Closure (BRAC) Agency, exception 1.5 of AFI 36-2601 applies, "Surveys requiring Office of Management and Budget (OMB) approval." The BRAC Agency gave AFCEE full authority to develop, administer, and improve BRAC's environmental project peer review program. The intent of the survey data is to improve the peer review program and is used within the scope of BRAC personnel employment for specific statistical purposes. Administering the survey to BRAC personnel falls under exception 3.17.11 of AFI 37-124, "Collections of information from Federal employees within the scope of their employment, unless the results are used for general statistical purposes."

Like the BRAC Agency employees, employees from the federal Environmental Protection Agency (EPA) fall under exception 1.5 of AFI 36-2601; since they are outside the Air Force, survey administration is controlled by AFI 37-124. Since the intent of the survey is to improve an environmental project peer review program, this is within the scope of employment EPA personnel. Exception 3.17.11 of AFI 37-124 exempts the survey from Office of Management and Budget approval.

The last two categories of peer reviewers who could be surveyed are state EPA employees and contractors. State EPA employees and contractors fall under exception 1.5 of AFI 36-2601 since administering surveys to these categories of people requires approval from the Office of Management and Budget (OMB) since they are outside the

Air Force. No exception within AFI 37-124 applies to these two categories, so OMB approval is required prior to conducting the survey on state EPA personnel and contractors. For planning factors, AFI 37-124, paragraph 3.10, recommends at least four months allowance for processing an OMB approval request.

Obtain Approvals as Required

Due to time constraints, there was not sufficient time to get the Office of Management and Budget approval--the approval was estimated to take at least four months. Accordingly, the survey was not sent outside the federal government to state employees or contractors. The following categories of people were surveyed: Air Force personnel at Brooks AFB, Wright-Patterson AFB, Air Education and Training Command, and Air Force Materiel Command; Environmental Protection Agency personnel; and Air Force Base Conversion Agency personnel.

Appendix H

Statistical Background

Multi-Item Variables

The predictor and criterion were multi-item variables. Multi-item variables areitems that area combination of two or more items. Without the aid of principle component analysis, the multi-item variables are simply an additive composite of the items they categorize (Hatcher, 1994:479).

Scales of Measurement

The variables are clearly classifying by scales of measurement. The four general scale of measurement categories are: nominal, ordinal, ratio, and interval. The nominal scale is the lowest level of measurement; this scale places the measured units of analysis into categories. The next highest level of measurement is ordinal. The ordinal scale ranks the units of analysis. With the interval scale, equal differences in the scale areequal numerically; however, the interval scale does not have an absolute zero. The ratio scale is the most descriptive measurement scale; equal differences in the scale areequal numerically *and* there is an absolute zero (Hatcher, 1994:7-9).

Quasi-Interval Variables

A further refinement in the categories of measurement is the quasi-interval scale. The Likert scale, used throughout the majority of the survey, does not have true equal-interval properties, such as those found on a Fahrenheit thermometer. Accordingly, survey items that rely on the Likert scale areoften categorized as quasi-interval (Hatcher, 1994:9).

Choosing the Correct Statistical Method

The important factors in choosing the correct statistic were the number and scale of both predictor and criterion variables. There were five quasi-interval and two nominal scale variables. There was one quasi-interval criterion variable. Either Analysis of Covariance (ANCOVA) or multiple linear regression can be used with this combination of variables. Multiple linear regression was the chosen statistical method because its key benefits meet the objectives of the analysis.

Multiple Linear Regression

Regression analysis has two key applications:

> (1) the prediction of values on a criterion variable based on acknowledge of values on predictor variables, and
> (2) the assessment of the relative degree to which each predictor variable accounts for variance in the criterion variable (Kachigan, 1991:186)

The primary application of regression analysis in this research was to assess the relative degree to which the seven predictor variables account for variance in the criterion effectiveness variable. With the knowledge that oral presentation accounts for 50 percent of the variance in peer review effectiveness, for example, focus would be placed on improving oral presentation both academically and through policy changes.

Multiple linear regression is a probabilistic model that relates the predictor variables to the criterion variable through an additive equation. The general multiple linear regression model is $Y = a + b_1x_1 + b_2x_2 + \ldots b_nx_n + \varepsilon$, where Y is the criterion variable, a is the intercept constant, b_n is the nonstandardized weight, x_n is the nth predictor variable, and ε is the error term. Further, it is assumed that the expected value of ε is zero, the variance of ε is σ^2, and ε is normally distributed (Devore, 1995:550; Hatcher, 1994:392).

Appendix I

Air Force Base Conversion Agency Peer Review Survey Comments

The Base Realignment and Closure Peer Reviews were held from 27 February to 6 March 1996. The projects under review are from the following installations: McClellan, Lowry, Myrtle Beach, Rickenbacker, Loring, Kelly, K.I. Sawyer, Castle, Wurtsmith, Mather, Plattsburg, Reese, March, Chanute, and Norton. 112 surveys were mailed and 41% were returned (N=46). The comments are reproduced as hand written with the exception of spelling errors. Information in brackets '[ ]' is added for clarity. Comments were received for the following four open-ended questions:

1. The following type of specialist should have been on the Peer Review Team:
2. The Peer Review Process would be improved by...
3. This survey would be improved by...
4. This survey should have asked...

The survey responses are grouped by the four questions below.

1. The following type of specialist should have been on the Peer Review Team:

- Team had [a] good mix of specialists.
- I do not know what specialists were on it. It was not discussed.
- Specialists familiar with cost effective technologies and related costs.
- Cost estimator.

- Our mix [was] O.K.
- SVE bioventing natural attenuation s/s.
- Mix was good. Maybe a cost analyst.
- Hydrologist, Geologist, Legal, Engineer, Chemist
- Right mix was used.
- People with industrial experience.
- Good coverage for our peer review; a little top heavy with toxicologists.

  EPA and State representatives from the region where the base is located. Other installations that have credible technical expertise to replace some of the AFCEE/San Antonio influence. Consider, for example, inviting some experienced base level RPM's to sit on the panel.
- Environmental economist.
- An experienced: cost estimator, value engineer (with environmental experience), and field chemist (with soil gas survey experience, for example).

The Peer Review Process would be improved by...

- Environmental economist.
- Stressing to the technical reviewers the major goals for BRAC facilities ... expediting clean-up/remediation to transfer property. Remind technical reviewers that the BRAC must work within regulatory requirements.

- More time, better prior preparation. Advanced materials to team. Make sure TCs [team chiefs] provide agenda/tech documents to PT [peer review team].
- Better coordination on time frames, meeting logistic and agenda preparation.
- Better coordination on time frames, meeting logistics and agenda preparation.
- Encourage peer reviewers to base their evaluation on scientific and risk to human health and not on ARARs [applicable or relevant and appropriate requirements] or regulator preference.
- Provide information to the Peer Review team 2 weeks ahead of time. Fully answer questions provided on peer review questionnaire. Provide clear detailed maps of each site. Provide peer review team detailed cost analysis.
- (1) Better meeting facilities for presentation of slides/overheads and audience participation. (2) Better preparation prior to the meeting (Peer Reviewers familiar with key documents and history and MAP etc.) (3) Better definition of purpose. Currently, seems like not enough time for peer reviewers to adequately understand a base's cleanup program to be able to provide valuable enough input to warrant the amount of staff time, reviewer time, regulator time spent jumping through all the hoops and preparation.
- Conducting it later in the process after RI's [remedial actions] are complete and FS [feasibility study] alternatives are ready to be evaluated (i.e., after sites are fully characterized.
- Intro needs to be given in an area where you can hear presenter. Terry Yonkers and Capt Smallwood speak softly and without emphasis on "what is important."

- Site visit(s).
- Having same reviewers as previous meetings so that part knowledge is carried through.
- Not attempting to review any compliance projects (600 series) but reviewing only restoration work.
- Reviewing the assumptions made on 1391's. Cost data should not be determined by Peer Review Team. Low dollar requirements, however, should not be the main focus of the review. Cost estimates should not be modified by Peer Review. The majority of time spent on Peer Review should not be on one or two low dollar items. The Program Manager and service agency should be allowed flexibility for some items that are deemed necessary due to local concerns. More weight should be given to non voting commentators who have spent considerable time on the projects and sites.
- (1) Providing clear, well organized descriptions of each project re: location, size, contamination (detailed backup available for reference), response action proposed, cost (detailed assumptions & estimates) and reason for the action, (2) Require (and pay) peer review team to spend adequate time (not less than 4 hrs.) to review the material before the review.
- Having the Base Program Manager prepared to fully engage in the Peer review process, rather than being defensive of his past program and decisions, etc.
- If the peer review team had been prepared. We found they were reading the package presented which took alot of time and their questions weren't always founded due to a quick look-through of the packages.

- The peer review team should be made to read the material before the meeting. The peer review should have some rules of thumb for evaluating cost estimates such as costs per square or cubic yard to evaluate areas or volumes to remediate.
- Getting 1/2 the amount of peer review participants out to the base for a week and evaluate the programs. It would be cheaper.
- longer lead time for peer review team to look over documentation, presentation of risk assessment details, and agenda.
- A better peer review would have occurred if the reviewers were given the pertinent documents, forms, etc., at least 1 week prior to the review, not the day of the review.
- The peer review had 3 chemists, 4 engineers, 3 hydrologists and 2 toxicologists. This seem to be a very wasteful use of technical resources for the USAF. A single representative from each discipline (scientific) should be sufficient either from AFCEE, contractor, or command.
- Provide more project descriptions & cost breakout.
- The project narratives and RACER estimates for the program were submitted on schedule, however, the info was not passed onto the peer review team in time for them to conduct any meaningful review.
- Sent data to AFCEE on first of month--Peer reviewers were given two notebooks (3") on morning of review.
- Adequately conveying to the base prior to the peer review what is to be expected. The base should be prepared to give a brief discussion of the status of the IRP program at the base, site locations, condition, etc. And then be prepared with budgeting and

technical presentations of the remedial alternatives selected for sites. Kelly AFB was not prepared. In addition, the presenters were so nervous/inexperienced that they were largely incoherent. The combination of lack of preparation and from communication skills on the part of the base negatively impacted the process of the KAFB peer review.

- Being realistic about what kind of technical review can be performed in one day. Thousands of team man-hours are spent contemplating RI/FS and RA documents in order to arrive at an appropriate remedial strategy. It is incredibly pretentious for some highly educated people to think they can validate so much team contemplation in one day.
- Ensuring that the prepared materials for the Peer Review were given to the peer review team prior to the meeting. In addition, the peer review should be timely as to when the facilities are in the position to make decisions on projects rather than just using the best available information at the time of the review.
- Availability of data from study contracts and sampling/analysis results to date applicable to each project.
- Peer review team having more time to review program/project [ ] prior to actual peer review. Provide more detail as to how the proposed remedial action(s) fits into the overall plan to close the site(s).
- Support from BRAC: let bases know what peer review is and tell them to cooperate like it or not!

- Using a facilitator who briefed ground rules and then held the team to these standards. The facilitator needed to keep the program on track. A larger room with a writing board to put up objectives for all to follow would have helped.
- 1) There were too many AFCEE people involved in the review. One AFCEE person added little value to the review and did crossword puzzles a large percentage of the time. 2) It seemed that the peer reviewers had not seen the programming documents ahead of the sessions--we had to take time at the beginning of each project presentation for the members to read the project narratives and cost estimates. 3) Better physical facilities would have helped. The hotel room we used was poorly set up for overhead projection and was too crowded. 4) More attention to the base request for presentation materials and aids. We had asked a couple of days ahead of time for dual-overhead projectors and screens. That didn't happen and we couldn't determine the request was even received or seriously considered. 5) Better communication of the process and the agenda.' We arrived in San Antonio the night before and it was not until an hour after the proposed start time the next morning before we were finally told where we'd be (which room in the hotel), and what the proposed schedule was. A lot of time was wasted the first morning just getting into the respective rooms, getting seated, etc.
- A) Improve the BECs' understanding of Chemical Sampling & Analysis (by providing/recommending training so they can answer or understand the following:

- For example, for analytical results, is it feasible the compounds detected were generated by AF operations (and if so, what were those operations?) or could they be lab contaminants (which would point to a cross-check by another lab) or are they naturally-occurring (and indicative of background concentrations)? Examples of questionable compounds are: 4-methyl-2-pentanone, 4-methyl-phenol, MEK, acetone, methylene chloride, beryllium, chloroform, arsenic, phthalates. (Also, the lst four compounds are readily biodegradable and thus great intrinsic remediation candidates.)

- How much confidence can one have in analytical results close to detection limits? [And how does one know what the detection limit is for each compound?] What is the significance of laboratory data qualifiers such as J (estimated value--concentration uncertain), U (compound not detected), B (analyte found in blank but not in sample)?

- For analytical costs, it would help for those who negotiate the scope of projects to understand what compounds are analyzed in the various EPA methods since sometimes compounds are analyzed redundantly or unnecessarily (i.e., inappropriately)--at greatly increased cost. Ex: EPA modified method 8015 (for TPH including JP-4, gasoline, etc.) combined with EPA 8020 (BTEX), EPA infrared method 418.1 (for grease & heavy petroleum residuals), a broad range of organics using EPA 8240 (GC/MS), target metals analyzed using EPA 60 1 0 vs. TAL (i.e., a wide range of) metals, etc.

- What is the threat to groundwater of various compounds detected in soil? For example, PAHs are practically insoluble in water, lead and most metals have very low solubility, etc.

- Also, with respect to PAHS, their detection (in the absence of other fuel components) in fuel spill locations indicates that biodegradation is taking place (i.e., that the lighter fuel fractions have already biodegraded or volatilized), and the PAHs will also biodegrade. (What is the biodegradation rate of PAHs--and other organics?) What is their W/W % in various fuels?) PAHs are very stable, bind tightly to the soil matrix, and are not a direct contact threat unless adsorbed to easily-windblown dust. What are typical PAH concentrations in the surface soil beside runways and heavily-traveled roadways? (i.e., Should we also be remediating this soil?)

- Once a compound has been labeled as a "contaminant" with regulator buy-in, it's very hard to get it unlabeled--so arguments need to be made early on against: 1) wrong or questionable analytical evidence and/or 2) faulty reasoning.

B) Require that cost estimates be summarized or attached to 1391s and narratives in the format of the HTRW RA WBS (Hazardous Toxic Radioactive Waste Remedial Action Work Breakdown Structure). This is the same estimate format (e.g., WBS) that RACER uses and would help standardize cost estimates and define their level of detail. Contractors should also be required to present their estimates in this format (to facilitate comparison to government estimates and compiling historical cost data). Also, a Historical Cost Database should be available for cost estimating and comparison purposes which holds all site description and cost data for past projects. This database should hold government estimates, negotiated costs, and final (completed project) cost data the HTRW RA WBS format.

C) The Peer Review Process would have definitely been more productive if the Michigan state regulator had been present.

3. This survey would be improved by:

- Providing it at the end (that day or 1 or 2 days after) of the review.
- Asking for a more specific description of the actual peer review.
- Adding a not applicable column and potentially an area for comments after each question.
- Getting it to the participant within one week of the peer review. Detailed memory fades 6 weeks later.
- Organizing and prerun to identify foci.
- Asking for comments after each of the 7 multiple-choice categories.
- OK
- Giving it immediately (or within 1 week) of the review--not 6-7 weeks later when my memory of this event has faded.

- Having participants complete it immediately after peer review. Kelly peer review was 6 weeks ago and only on of several peer reviews.
- Good survey! Could have allowed for remarks under negative relies but at the end is ok.
- No recommendations.
- The survey is adequate.
- Adding comment fields to questions.
- More questions related to the value added by such peer reviews.
- Asking more detailed, specific questions regarding the behavior and focus of the peer review team, and the peer review facilities. Also, clarify questions; what is the difference between question 48 and 49 on this survey?
- OK
- Make entire survey more focused on the particular parties you want the feedback from. Perhaps one questionnaire for the Peer Review Panel members to feedback on the process and the base presenters, and another questionnaire for the base-level folks that came and presented. You're trying to make one questionnaire do too much and cover too many different players and points of-view.

4. This survey should have asked...

- Why peer reviewers from AFCEE itself did not talk.
- ok.

- Is there a true need to conduct peer reviews? At what point in the clean up process should peer reviews be conducted? Who should set policies on how to conduct peer reviews?
- Were accommodations OK: Room, business support, temp, etc. Need a new location next year.
- Supplied more room for comments after each question.
- Good survey!
- Sooner.
- Sufficient.
- What phase is your program in?
- What happened at the peer review. I felt the process was good and peer reviewers played their roles well. There was some antagonism between some of the peer reviewers with each other and the base. The base was not given proper description of the AFBCA peer review process and were not as well prepared as they could have been. Also, due to a change in schedule, generated by AFBCA, Kelly's peer review schedule had to be severely altered. This did not allow Kelly to present their program in a way that was easiest for them.
- # 28 is not a question and there was no section to rate the facilitator, the facilities, or the feedback to the BEC.
- More about the location and pre-planning for the peer review sessions, more about physical facilities and layout, and more feedback on options to do peer review via

VTCN teleconference. Also about how the feeling to have the peer review team travel to the base rather than have the base travel to peer review. Being on-site to do the peer review might add more to the understanding of the program needs, community situation, etc.And the survey should ask if those completing the questionnaires would like to know the results of the survey and what changes in the process came about as a result.

Appendix J

Air Education & Training Command Survey Comments

The Air Education and Training Command conducted an environmental remediation project peer review from 19 April 1996 to 10 May 1996. The following Air Force bases presented projects for review: Vance, Reese, Laughlin, Keesler, Lackland, Randolf, Tyndall, and Maxwell. The following responses are word-for-word, with the exception of spelling errors, from the open-ended survey questions. Surveys were mailed under letter from HQ AETC to each base Remedial Program Manager (RPM) on 1 July 1996. Eight surveys were mailed and returned completed. Seven of the eight RPMs completed the open-ended questions. Information in brackets "[ ]" is added for clarity. Comments were received for the following four open-ended questions:

1. The following type of specialist should have been on the Peer Review Team:
2. The Peer Review Process would be improved by...
3. This survey would be improved by...
4. This survey should have asked...

The survey responses are grouped by the four questions below.

1. The following type of specialist should have been on the Peer Review Team:

- Technology experts.
- Risk assessors, lawyer, field remediation construction person.

- Team represented a good mix.
- Geologist.

2. The Peer Review Process would be improved by...

- Going back to having experts review processes before we get to writing narratives, 1391's, etc. Like the DNAPL [Dense Non-Acquious Phase Liquids] forum. We wasted a lot of time in program submittal. This should have been decided long before. It is too labor intensive to spend over 160 hours working on a program submittal and then trashing over half of it.
- More clear-cut understanding of what was to be expected during review and a "minutes" of the review afterwards regarding suggestions provided.
- (1) Having the above specialists on the team. (2) Require photos of site in Question! (3) With more and more sites being remediated, look at other projects that may be similar to those being reviewed.
- Having other non MAJCOM personnel i.e. regulators, AFCEE, etc. participate on the Peer Review Team to offer additional insight, experience, etc.
- Having a geologist present during peer review would only benefit the entire process.
- RPM providing to the peer review team the review documents sooner (more time to review in advance.)
- Reduce numbers on review team; publish more concise peer review guidance; publish better review agenda.

3. This survey would be improved by...

- If it applied more to AETC
- Branch this out into several sections: As Base RPM looking at Peer Review, as HQ RPM looking at Base RPM & Peer Review and Peer Review looking at HQ & Base RPM. Have east section answer sections which pertain, and then disseminate results to everyone.
- Provide it earlier (i.e. right after the Peer Review has been conducted).

4. This survey should have asked...

- (1) What was liked/disliked about the peer review. (2) Do you think your projects were fairly reviewed?
- "Were there any areas the peer review team did not take into consideration?"

Appendix K

Air Force Materiel Command Survey Comments

The Air Force Materiel Command conducted an environmental remediation project peer review from 20 August 1996 through 22 August 1996. The following bases presented projects for review: Robins, Eglin, Kirtland, Hill, and Edwards. Surveys were administered at the end of each peer review session. 21 surveys were administered to peer review participants; the response rate was 76% (N=16). The comments are reproduced as hand written with the exception of spelling errors. Information in brackets '[ ]' is added for clarity. Comments were received for the following four open-ended questions:

1. The following type of specialist should have been on the Peer Review Team:
2. The Peer Review Process would be improved by...
3. This survey would be improved by...
4. This survey should have asked...

The survey responses are grouped by the four questions below.

1. The following type of specialist should have been on the Peer Review Team:

- Engineers, scientists, and ecologists should be on the team. There were no ecologists.
- Risk assessor.
- Biologist or Limnologist.
- Design/construction costs?

- Chemical Engineer.
- Team was complete.
- Air sparging.
- The mix was fine.

2. The Peer Review Process would be improved by...

- It would be improved by handling it like academic journal peer review instead of judge & jury.
- Overheads & list of conclusions on recommendations.
- Complete project packages provided early enough in advance to ensure adequate reading & reflection.
- Agenda, complete documentation of projects (map).
- I believe peer review process is appropriate and necessary. I would not change anything. It is well run and the peer review team is very knowledgeable.
- The peer review team sending a list of questions or comments to the base program manager prior to the peer review so that time would be available to research specific issues concerning the project.
- Provide some advance "warning" of questions; changing format of document (peer summary report), very labor intensive to create/info already available in different formats.

- Better definition of the type of information that is important to the decision process for the peer review.
- Having regulators on this team so their point of view or feelings can be given directly to peer reviewers. Also the team should have people familiar with the political climate & regulations of the base under review. Peer review should concentrate on the technology and not the money. Basically a waste of time and money.
- Taking into account that the base is under pressure to institute cleanups and that the base staff & contractors are competent.
- By having handouts for everyone.
- Clearly defining the objectives of all USAF Peer Reviews--is it to review and advise on best technologies? Or to do Value Engineering? Or to see if the "politically correct" alternative is chosen?
- I would like to see a written follow-up from the committee.

3. This survey would be improved by...

- Less questions.
- Shorter!
- Shorter.
- Shorten.
- Good questions!
- Make it much shorter. One page only.

- N/A.
- Ask if the peer review provided any meaningful input. Did the (any) projects change or were improved by Peer Review?
- The form is not clear on whether I should respond to each category i.e. Q 33-40 Look like they should only be responded to by the peer review members.

4. This survey should have asked...

- Offer money for completing survey.
- How did you "feel" about the experience?
- ?
- About the preparation time. Preparing documents is a lot of work + time. The level of effort required should be communicated from the peer review team to the project managers.
- N/A.
- Project type being reviewed (i.e. what contam., media, and remedy) what changes were suggested, then your study can show what types of remediation need peer review, and which types are a waste of time, TDY travel and staff time. I believe the peer review should not be convened for presumptive remedies since they have already gone through regulatory and public review and have been identified as the best technology for the contamination/media mix.

Appendix L

Air Force Materiel Command Peer Review Observations

Background

The Air Force Materiel Command (AFMC) held five environmental remediation project peer reviews sequentially from 20 August 1996 through 22 August 1996. The Robins Air Force Base (AFB) peer review was held from 0800 to 1200 on 20 August 1996. The Eglin AFB peer review was held from 1300 to 1700 on 20 August and was completed the following morning. The Kirtland AFB peer review was held from 1300 to 1500 on 21 August. The Hill AFB peer review started at 0800 on 22 August and finished around 1000 hours. The Edwards AFB peer review was held from 1300 to 1700 on 22 August 1996.

The AFMC environmental remediation peer review team leader--a geologist--facilitated each of the peer reviews. There were four other members on the technical peer review team: two AFIT professors (PhDs)--both environmental engineers--and two base project managers (PMs)--both environmental engineers. The two professors brought academic, managerial, and dated field experience to the table, while the two PMs added a current, practical field experience perspective. The number of PMs representing each of the five bases varied from a minimum of one PM representing Kirland AFB to a maximum of four PMs and contractors representing Robins AFB, Eglin AFB, and Hill AFB. Edwards AFB had three PMs.

The peer reviews were held in a conference room that comfortably seats approximately eight at the table and three at the back. There was an overhead and easel with writing tablet available for presentations. There were no refreshments available. The room was on the warm-side of comfortable at times.

Robins AFB Peer Review, Tuesday, 20 August 1996

The Duck Lake site was the first project reviewed. There was no overhead or overview of the overall Robins program nor the specific project. However, it was apparent that the peer review team had previously visited the site. The peer review team leader suggested the team "cut through the chase and start firing questions." The lead PM wanted to give an "advocacy briefing." The peer review team began by focusing on the assumptions and case studies used for determining ecological risk. Great Lakes studies were discussed and their relevance to the Duck Lake project was assessed. Insitu natural attenuation was proposed and assessed for a cost effective alternative.

The peer review team each received two written project packages about 20 pages in length prior to the review. The PM brought two project binders with additional detailed information for the review.

The peer review team began jumping around with questions in different directions. The peer review team leader attempted to focus the review stating, "We keep jumping around." Most analysis was focused on the risk drivers and assumptions. Following several hours of discussion, the team leader requested consensus that, "Something needs to be done." The next questions are, how and to what levels?

The peer review team leader proposed two alternatives and then reached consensus. The team leader summarized recommendations, built consensus, and guided the process.

The next project evaluated was Landfill 3. The peer review team leader suggested natural attenuation, then went on to say that we will assume regulators will say “no” to natural attenuation. The peer review team was against the use of slurry walls.

The peer review team pre-approved the project up to a cost of $720,000. If above that cost, the project will be reevaluated.

Robins AFB is currently using wicks to collect small amounts of JP4 product from wells around a JP4 spill site. The peer review team recommended discontinuing use of the wicks and to look into bioventing since it is a presumptive remedy for JP4 spills. The peer review team leader said, “We always ‘buy-in’ to bioventing for JP4.”

Eglin AFB Peer Review, Tuesday-Wednesday, 20-21 August 1996

Like the Robins AFB peer review, the Eglin peer review did not utilize the overhead; but, an overview of the projects under review was given. A site map overhead would have improved the effectiveness of the overview. The peer review team leader stated, after the overview, “Lets focus on one thing at a time.” The leader later refocused the review, “Lets ‘sorta’ stay focused here.” As the review dived into detail the leader stated, “We can’t spend too much time on this, we have a lot of projects to get through.” The leader stated later, “Lets not jump around here. What are our recommendations?”

One of the PMs asked, “What are you looking for in a PSR [project summary report]? The peer review members replied, “Exactly what are you proposing to do? Be consistent." "What do you have? What are you proposing to do? How will it help?" "We need maps." The peer review team leader added, "Risk and regulatory framework; facts and figures; and articulation of project. Focus on the discussion section and strategy. If the project is marginal, a note from your lawyer may be helpful."

A peer review team member stated, "Peer review should be technical review, not a review for funding." Upper management should make the non-technical decisions. One of the PMs stated, that he feels caught in the middle because he cannot bargain [with the regulators] due to required peer review. Another peer review team member stated, "It is certainly not the goal of any of us to cut things out of projects."

Kirtland AFB Peer Review, Wednesday, 21 August 1996

The PM brought color pictures of site that were very helpful. The "table-top" briefing was very thorough. Did not use overhead, but color photos were excellent substitute. The driver for the project was the public, not technical. Current peer review process requires review team to agree with necessity of project on technical grounds. Discussion added that this was a "Catch-22" because the non-technical rationale was valid but the technical rationale was lacking. Further discussion added that it makes more sense for the peer review to review on a technical basis and management to approve due to public, regulatory, or political drivers.

Hill AFB Peer Review, Thursday, 22 August 1996

An overhead used with map and overview briefing. The review went quickly and smoothly.

Discussion followed Hill AFB Peer Review

Peer review team should be technical. Scientists should not be asked to make political and management decisions. Focus should be on technical issues not funding.

The funding measure of merit is flawed. Form 1391s are rough estimates and can be easily inflated. Refined cost information does not indicate a good technical peer review. Technical costs may increase costs from 1391s. Increase cost but better project. This would result in a poor rating for the peer review with the funding measure of merit. 1391s do not capture long-term O&M. Technical peer review may change technical parameters and decrease O&M but increase 1391 costs. Overall cost may decrease but not be captured on 1391. The peer review team must have practical or lab experience not just academic. Not all costs on 1391s need to be approved. Peer review team should look for costs that are noticeably high or low for review.

Edwards AFB Peer Review, Thursday, 22 August 1996

Peer review team leader gave ground rules before review began: first we will review whether the action is needed or justified; then, we will review the technology and costs. The PM used an overhead with color maps; the briefing was very effective. The peer review team leader explained that the peer review team has become more of a technical review, not a review based on "other factors." The peer review team recommendation is no longer required for funding. One peer review team member stated, "We recognize there are other issues such as regulators and public; these issues must be justified by PMs, not on a technical basis." The PM stated that she would like to see if what they [now] planned to do met the peer review recommendations.

Analysis

The following results were drawn from observing the Air Force Materiel Command peer reviews:

1. Peer reviews would be more effective if they reviewed technical issues and made recommendations on a technical basis to management. Management is better suited to pulse the regulatory, public, and political climate and make "management" decisions. Scientists and engineers are better suited to make "technical decisions."
2. Peer reviews with PMs that utilized overhead maps or photos to explain sites appeared to be more effective and more efficient. A introductory briefing of the "big picture" was also beneficial when given. The quality of the oral presentation was critical to the peer review team understanding the projects under review.
3. The peer review team was heavily slanted toward environmental engineering. The team could be better balanced with a risk assessor, an ecologist, a chemist, and possibly a cost analyst. These specialists would lend additional credibility and specific expertise to the team.
4. The process would flow more smoothly if the ground rules were explicitly spelled out. Recommend explicitly assigning the role of facilitator. Also, an agenda would be very

helpful to keep the peer review focused. Agenda items could be developed from areas that peer review team members need further explanation. This would also facilitate PM preparation. Misunderstandings developed because PMs did not clearly know what was expected of them.

Appendix M

Survey Item Frequency Distributions

The SAS® output that follows contains frequency distributions generated from the peer review process survey data. The frequency distributions are useful because they pictorally capture the trends and shape of the distributions of the data.

```
The SAS System                                   91
                                                   12:11 Wednesday,
August 28, 1996

                                 Cumulative  Cumulative
EFF1   Frequency     Percent     Frequency     Percent
-------------------------------------------------------
  1          17       25.0            17        25.0
  2          24       35.3            41        60.3
  3          15       22.1            56        82.4
  4          11       16.2            67        98.5
  5           1        1.5            68       100.0

             Frequency Missing = 2
```

```
                      The SAS System
92
                                                   12:11 Wednesday,
August 28, 1996

                                 Cumulative  Cumulative
EFF2   Frequency     Percent     Frequency     Percent
-------------------------------------------------------
  1          17       25.0            17        25.0
  2          33       48.5            50        73.5
  3          13       19.1            63        92.6
  4           5        7.4            68       100.0

             Frequency Missing = 2
```

```
                      The SAS System
93
                                                   12:11 Wednesday,
August 28, 1996

                                 Cumulative  Cumulative
EFF3   Frequency     Percent     Frequency     Percent
-------------------------------------------------------
  1          12       17.9            12        17.9
  2          20       29.9            32        47.8
  3          25       37.3            57        85.1
  4           9       13.4            66        98.5
  5           1        1.5            67       100.0

             Frequency Missing = 3
```

```
                      The SAS System
94
                                                   12:11 Wednesday,
August 28, 1996

                                 Cumulative  Cumulative
EFF4   Frequency     Percent     Frequency     Percent
-------------------------------------------------------
  1          19       27.9            19        27.9
  2          35       51.5            54        79.4
  3           7       10.3            61        89.7
  4           6        8.8            67        98.5
  5           1        1.5            68       100.0

             Frequency Missing = 2
```

```
                      The SAS System
95
                                                   12:11 Wednesday,
August 28, 1996

                                 Cumulative  Cumulative
EFF5   Frequency     Percent     Frequency     Percent
-------------------------------------------------------
  1          14       20.6            14        20.6
  2          23       33.8            37        54.4
  3          21       30.9            58        85.3
  4           9       13.2            67        98.5
  5           1        1.5            68       100.0

             Frequency Missing = 2
```

```
                      The SAS System
96
                                                   12:11 Wednesday,
August 28, 1996

                                 Cumulative  Cumulative
EFF6   Frequency     Percent     Frequency     Percent
-------------------------------------------------------
  1          13       19.1            13        19.1
  2          34       50.0            47        69.1
  3          16       23.5            63        92.6
  4           2        2.9            65        95.6
  5           3        4.4            68       100.0

             Frequency Missing = 2
```

```
                      The SAS System
97
                                                   12:11 Wednesday,
August 28, 1996

                                 Cumulative  Cumulative
EFF7   Frequency     Percent     Frequency     Percent
-------------------------------------------------------
  1           9       13.2             9        13.2
  2          25       36.8            34        50.0
  3          22       32.4            56        82.4
  4           8       11.8            64        94.1
  5           4        5.9            68       100.0

             Frequency Missing = 2
```

```
                      The SAS System
98
                                                   12:11 Wednesday,
August 28, 1996

                                 Cumulative  Cumulative
EFF8   Frequency     Percent     Frequency     Percent
-------------------------------------------------------
  1          15       22.1            15        22.1
  2          36       52.9            51        75.0
  3           9       13.2            60        88.2
  4           7       10.3            67        98.5
  5           1        1.5            68       100.0

             Frequency Missing = 2
```

```
                      The SAS System
99
                                                   12:11 Wednesday,
August 28, 1996

                                 Cumulative  Cumulative
EFF9   Frequency     Percent     Frequency     Percent
-------------------------------------------------------
  1           8       12.1             8        12.1
  2          20       30.3            28        42.4
  3          26       39.4            54        81.8
  4          11       16.7            65        98.5
  5           1        1.5            66       100.0

             Frequency Missing = 4
```

```
                      The SAS System
100
                                                   12:11 Wednesday,
August 28, 1996

                                  Cumulative  Cumulative
EFF10   Frequency     Percent     Frequency     Percent
--------------------------------------------------------
   1           8       11.8             8        11.8
   2          40       58.8            48        70.6
   3          10       14.7            58        85.3
   4           8       11.8            66        97.1
   5           2        2.9            68       100.0

              Frequency Missing = 2
```

```
                      The SAS System
101
                                                   12:11 Wednesday,
August 28, 1996

                                  Cumulative  Cumulative
EFF11   Frequency     Percent     Frequency     Percent
--------------------------------------------------------
   1          16       23.9            16        23.9
   2          30       44.8            46        68.7
   3          11       16.4            57        85.1
   4           6        9.0            63        94.0
   5           4        6.0            67       100.0

              Frequency Missing = 3
```

```
                              The SAS System
102
                                                  12:11 Wednesday,
August 28, 1996

                                           Cumulative  Cumulative
             EFF12   Frequency   Percent    Frequency    Percent
             ------------------------------------------------------
                 1         23      33.8           23       33.8
                 2         21      30.9           44       64.7
                 3          7      10.3           51       75.0
                 4         12      17.6           63       92.6
                 5          5       7.4           68      100.0

                          Frequency Missing = 2


                              The SAS System
103
                                                  12:11 Wednesday,
August 28, 1996

                                           Cumulative  Cumulative
              FOC1   Frequency   Percent    Frequency    Percent
             ------------------------------------------------------
                 1         12      17.6           12       17.6
                 2         30      44.1           42       61.8
                 3         17      25.0           59       86.8
                 4          8      11.8           67       98.5
                 5          1       1.5           68      100.0

                          Frequency Missing = 2


                              The SAS System
104
                                                  12:11 Wednesday,
August 28, 1996

                                           Cumulative  Cumulative
              FOC2   Frequency   Percent    Frequency    Percent
             ------------------------------------------------------
                 1         10      14.7           10       14.7
                 2         31      45.6           41       60.3
                 3         16      23.5           57       83.8
                 4         10      14.7           67       98.5
                 5          1       1.5           68      100.0

                          Frequency Missing = 2


                              The SAS System
105
                                                  12:11 Wednesday,
August 28, 1996

                                           Cumulative  Cumulative
              FOC3   Frequency   Percent    Frequency    Percent
             ------------------------------------------------------
                 1          7      10.3            7       10.3
                 2         29      42.6           36       52.9
                 3         18      26.5           54       79.4
                 4         13      19.1           67       98.5
                 5          1       1.5           68      100.0

                          Frequency Missing = 2


                              The SAS System
106
                                                  12:11 Wednesday,
August 28, 1996

                                           Cumulative  Cumulative
              FOC4   Frequency   Percent    Frequency    Percent
             ------------------------------------------------------
                 1         14      20.6           14       20.6
                 2         29      42.6           43       63.2
                 3         16      23.5           59       86.8
                 4          8      11.8           67       98.5
                 5          1       1.5           68      100.0

                          Frequency Missing = 2


                              The SAS System
107
                                                  12:11 Wednesday,
August 28, 1996

                                           Cumulative  Cumulative
              FOC5   Frequency   Percent    Frequency    Percent
             ------------------------------------------------------
                 1         10      14.7           10       14.7
                 2         37      54.4           47       69.1
                 3         13      19.1           60       88.2
                 4          8      11.8           68      100.0

                          Frequency Missing = 2


                              The SAS System
108
                                                  12:11 Wednesday,
August 28, 1996

                                           Cumulative  Cumulative
              FOC6   Frequency   Percent    Frequency    Percent
             ------------------------------------------------------
                 1          9      13.4            9       13.4
                 2         38      56.7           47       70.1
                 3         14      20.9           61       91.0
                 4          6       9.0           67      100.0

                          Frequency Missing = 3
```

```
                              The SAS System
109
                                                  12:11 Wednesday,
August 28, 1996

                                           Cumulative  Cumulative
              FOC7   Frequency   Percent    Frequency    Percent
             ------------------------------------------------------
                 1          9      13.4            9       13.4
                 2         25      37.3           34       50.7
                 3         20      29.9           54       80.6
                 4         10      14.9           64       95.5
                 5          3       4.5           67      100.0

                          Frequency Missing = 3


                              The SAS System
110
                                                  12:11 Wednesday,
August 28, 1996

                                           Cumulative  Cumulative
              FOC8   Frequency   Percent    Frequency    Percent
             ------------------------------------------------------
                 1          6       8.8            6        8.8
                 2         25      36.8           31       45.6
                 3         20      29.4           51       75.0
                 4         17      25.0           68      100.0

                          Frequency Missing = 2


                              The SAS System
111
                                                  12:11 Wednesday,
August 28, 1996

                                           Cumulative  Cumulative
              FOC9   Frequency   Percent    Frequency    Percent
             ------------------------------------------------------
                 1          3       4.4            3        4.4
                 2         15      22.1           18       26.5
                 3         32      47.1           50       73.5
                 4         15      22.1           65       95.6
                 5          3       4.4           68      100.0

                          Frequency Missing = 2


                              The SAS System
112
                                                  12:11 Wednesday,
August 28, 1996

                                           Cumulative  Cumulative
             FOC10   Frequency   Percent    Frequency    Percent
             ------------------------------------------------------
                 1         19      28.4           19       28.4
                 2         33      49.3           52       77.6
                 3         10      14.9           62       92.5
                 4          3       4.5           65       97.0
                 5          2       3.0           67      100.0

                          Frequency Missing = 3


                              The SAS System
113
                                                  12:11 Wednesday,
August 28, 1996

                                           Cumulative  Cumulative
             FOC11   Frequency   Percent    Frequency    Percent
             ------------------------------------------------------
                 1         10      14.7           10       14.7
                 2         27      39.7           37       54.4
                 3         20      29.4           57       83.8
                 4         11      16.2           68      100.0

                          Frequency Missing = 2


                              The SAS System
114
                                                  12:11 Wednesday,
August 28, 1996

                                           Cumulative  Cumulative
             FOC12   Frequency   Percent    Frequency    Percent
             ------------------------------------------------------
                 1          7      10.6            7       10.6
                 2         17      25.8           24       36.4
                 3         31      47.0           55       83.3
                 4         10      15.2           65       98.5
                 5          1       1.5           66      100.0

                          Frequency Missing = 4


                              The SAS System
115
                                                  12:11 Wednesday,
August 28, 1996

                                           Cumulative  Cumulative
              AGN1   Frequency   Percent    Frequency    Percent
             ------------------------------------------------------
                 1          5      10.6            5       10.6
                 2         27      57.4           32       68.1
                 3         13      27.7           45       95.7
                 4          2       4.3           47      100.0

                          Frequency Missing = 23


                              The SAS System
116
```

```
                                       12:11 Wednesday,
August 28, 1996

                                                Cumulative  Cumulative
              AGN2   Frequency    Percent     Frequency     Percent
              ---------------------------------------------------------
                 1          7       14.6             7        14.6
                 2         26       54.2            33        68.8
                 3         11       22.9            44        91.7
                 4          3        6.3            47        97.9
                 5          1        2.1            48       100.0

                          Frequency Missing = 22


                              The SAS System
117
                                       12:11 Wednesday,
August 28, 1996

                                                Cumulative  Cumulative
              AGN3   Frequency    Percent     Frequency     Percent
              ---------------------------------------------------------
                 1         11       22.9            11        22.9
                 2         22       45.8            33        68.8
                 3          8       16.7            41        85.4
                 4          6       12.5            47        97.9
                 5          1        2.1            48       100.0

                          Frequency Missing = 22


                              The SAS System
118
                                       12:11 Wednesday,
August 28, 1996

                                                Cumulative  Cumulative
              AGN4   Frequency    Percent     Frequency     Percent
              ---------------------------------------------------------
                 1          7       16.3             7        16.3
                 2         31       72.1            38        88.4
                 3          4        9.3            42        97.7
                 5          1        2.3            43       100.0

                          Frequency Missing = 27


                              The SAS System
119
                                       12:11 Wednesday,
August 28, 1996

                                                Cumulative  Cumulative
              WPR1   Frequency    Percent     Frequency     Percent
              ---------------------------------------------------------
                 1          6        9.1             6         9.1
                 2         26       39.4            32        48.5
                 3         15       22.7            47        71.2
                 4         14       21.2            61        92.4
                 5          5        7.6            66       100.0

                          Frequency Missing = 4


                              The SAS System
120
                                       12:11 Wednesday,
August 28, 1996

                                                Cumulative  Cumulative
              WPR2   Frequency    Percent     Frequency     Percent
              ---------------------------------------------------------
                 1          6        9.1             6         9.1
                 2         21       31.8            27        40.9
                 3         15       22.7            42        63.6
                 4         18       27.3            60        90.9
                 5          6        9.1            66       100.0

                          Frequency Missing = 4


                              The SAS System
121
                                       12:11 Wednesday,
August 28, 1996

                                                Cumulative  Cumulative
              WPR3   Frequency    Percent     Frequency     Percent
              ---------------------------------------------------------
                 1         12       17.9            12        17.9
                 2         33       49.3            45        67.2
                 3         13       19.4            58        86.6
                 4          6        9.0            64        95.5
                 5          3        4.5            67       100.0

                          Frequency Missing = 3


                              The SAS System
122
                                       12:11 Wednesday,
August 28, 1996

                                                Cumulative  Cumulative
              WPR4   Frequency    Percent     Frequency     Percent
              ---------------------------------------------------------
                 1          9       13.4             9        13.4
                 2         30       44.8            39        58.2
                 3         16       23.9            55        82.1
                 4         10       14.9            65        97.0
                 5          2        3.0            67       100.0

                          Frequency Missing = 3


                              The SAS System
123
```

```
                                       12:11 Wednesday,
August 28, 1996

                                                Cumulative  Cumulative
              OPR1   Frequency    Percent     Frequency     Percent
              ---------------------------------------------------------
                 1         15       22.1            15        22.1
                 2         31       45.6            46        67.6
                 3         14       20.6            60        88.2
                 4          6        8.8            66        97.1
                 5          2        2.9            68       100.0

                          Frequency Missing = 2


                              The SAS System
124
                                       12:11 Wednesday,
August 28, 1996

                                                Cumulative  Cumulative
              OPR2   Frequency    Percent     Frequency     Percent
              ---------------------------------------------------------
                 1         14       20.6            14        20.6
                 2         28       41.2            42        61.8
                 3         12       17.6            54        79.4
                 4         12       17.6            66        97.1
                 5          2        2.9            68       100.0

                          Frequency Missing = 2
```

```
125                              The SAS System
                                                   12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
              OPR3   Frequency   Percent   Frequency    Percent
              ---------------------------------------------------
                 1         18      26.5          18       26.5
                 2         38      55.9          56       82.4
                 3          7      10.3          63       92.6
                 4          3       4.4          66       97.1
                 5          2       2.9          68      100.0

                            Frequency Missing = 2
```

```
126                              The SAS System
                                                   12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
              OPR4   Frequency   Percent   Frequency    Percent
              ---------------------------------------------------
                 1         17      25.0          17       25.0
                 2         34      50.0          51       75.0
                 3         12      17.6          63       92.6
                 4          3       4.4          66       97.1
                 5          2       2.9          68      100.0

                            Frequency Missing = 2
```

```
127                              The SAS System
                                                   12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
              OPR5   Frequency   Percent   Frequency    Percent
              ---------------------------------------------------
                 1         19      27.9          19       27.9
                 2         36      52.9          55       80.9
                 3          9      13.2          64       94.1
                 4          3       4.4          67       98.5
                 5          1       1.5          68      100.0

                            Frequency Missing = 2
```

```
128                              The SAS System
                                                   12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
              OPR6   Frequency   Percent   Frequency    Percent
              ---------------------------------------------------
                 1         13      19.1          13       19.1
                 2         18      26.5          31       45.6
                 3         24      35.3          55       80.9
                 4         10      14.7          65       95.6
                 5          3       4.4          68      100.0

                            Frequency Missing = 2
```

```
129                              The SAS System
                                                   12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
              OPR7   Frequency   Percent   Frequency    Percent
              ---------------------------------------------------
                 1          1       1.5           1        1.5
                 2          6       9.0           7       10.4
                 3         43      64.2          50       74.6
                 4         14      20.9          64       95.5
                 5          3       4.5          67      100.0

                            Frequency Missing = 3
```

```
130                              The SAS System
                                                   12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
              OPR8   Frequency   Percent   Frequency    Percent
              ---------------------------------------------------
                 2          7      10.4           7       10.4
                 3         50      74.6          57       85.1
                 4          7      10.4          64       95.5
                 5          3       4.5          67      100.0

                            Frequency Missing = 3
```

```
131                              The SAS System
                                                   12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
              TMC1   Frequency   Percent   Frequency    Percent
              ---------------------------------------------------
                 1          3       4.5           3        4.5
                 2          7      10.4          10       14.9
                 3         53      79.1          63       94.0
                 4          3       4.5          66       98.5
                 5          1       1.5          67      100.0

                            Frequency Missing = 3
```

```
132                              The SAS System
                                                   12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
              TMC2   Frequency   Percent   Frequency    Percent
              ---------------------------------------------------
                 1          2       2.9           2        2.9
                 2          8      11.8          10       14.7
                 3         52      76.5          62       91.2
                 4          6       8.8          68      100.0

                            Frequency Missing = 2
```

```
133                              The SAS System
                                                   12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
              TMC3   Frequency   Percent   Frequency    Percent
              ---------------------------------------------------
                 1         10      14.7          10       14.7
                 2         42      61.8          52       76.5
                 3         11      16.2          63       92.6
                 4          4       5.9          67       98.5
                 5          1       1.5          68      100.0

                            Frequency Missing = 2
```

```
134                              The SAS System
                                                   12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
              TMC4   Frequency   Percent   Frequency    Percent
              ---------------------------------------------------
                 1         12      17.6          12       17.6
                 2         38      55.9          50       73.5
                 3         15      22.1          65       95.6
                 4          3       4.4          68      100.0

                            Frequency Missing = 2
```

```
135                              The SAS System
                                                   12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
              TMC5   Frequency   Percent   Frequency    Percent
              ---------------------------------------------------
                 1          5       7.4           5        7.4
                 2         23      33.8          28       41.2
                 3         25      36.8          53       77.9
                 4         13      19.1          66       97.1
                 5          2       2.9          68      100.0

                            Frequency Missing = 2
```

```
136                              The SAS System
                                                   12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
              TMC6   Frequency   Percent   Frequency    Percent
              ---------------------------------------------------
                 1          8      11.9           8       11.9
                 2         31      46.3          39       58.2
                 3         22      32.8          61       91.0
                 4          6       9.0          67      100.0

                            Frequency Missing = 3
```

```
137                              The SAS System
                                                   12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
              RVC1   Frequency   Percent   Frequency    Percent
              ---------------------------------------------------
                 1          2       2.9           2        2.9
                 2         14      20.6          16       23.5
                 3         21      30.9          37       54.4
                 4         20      29.4          57       83.8
                 5         11      16.2          68      100.0

                            Frequency Missing = 2
```

```
138                              The SAS System
                                                   12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
              RVC2   Frequency   Percent   Frequency    Percent
              ---------------------------------------------------
                 1         15      22.1          15       22.1
                 2         36      52.9          51       75.0
                 3          9      13.2          60       88.2
                 4          7      10.3          67       98.5
                 5          1       1.5          68      100.0

                            Frequency Missing = 2
```

```
139                              The SAS System
```

```
                                                 12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
         RVC3   Frequency     Percent      Frequency    Percent
         -------------------------------------------------------
            1          11        16.2             11       16.2
            2          29        42.6             40       58.8
            3          15        22.1             55       80.9
            4          11        16.2             66       97.1
            5           2         2.9             68      100.0

                         Frequency Missing = 2
```

```
                              The SAS System
140
                                                 12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
         RVC4   Frequency     Percent      Frequency    Percent
         -------------------------------------------------------
            1          12        17.6             12       17.6
            2          38        55.9             50       73.5
            3          13        19.1             63       92.6
            4           4         5.9             67       98.5
            5           1         1.5             68      100.0

                         Frequency Missing = 2
```

```
                              The SAS System
141
                                                 12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
         RVC5   Frequency     Percent      Frequency    Percent
         -------------------------------------------------------
            1          11        16.2             11       16.2
            2          35        51.5             46       67.6
            3          17        25.0             63       92.6
            4           5         7.4             68      100.0

                         Frequency Missing = 2
```

```
                              The SAS System
142
                                                 12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
         RVC6   Frequency     Percent      Frequency    Percent
         -------------------------------------------------------
            1          10        14.7             10       14.7
            2          38        55.9             48       70.6
            3          16        23.5             64       94.1
            4           3         4.4             67       98.5
            5           1         1.5             68      100.0

                         Frequency Missing = 2
```

```
                              The SAS System
143
                                                 12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
         JBI1   Frequency     Percent      Frequency    Percent
         -------------------------------------------------------
            0           8        11.8              8       11.8
            3           8        11.8             16       23.5
            4          21        30.9             37       54.4
            5          18        26.5             55       80.9
            9          13        19.1             68      100.0

                         Frequency Missing = 2
```

```
                              The SAS System
144
                                                 12:11 Wednesday,
August 28, 1996

                                          Cumulative  Cumulative
         JBI2   Frequency     Percent      Frequency    Percent
         -------------------------------------------------------
            0           1         1.5              1        1.5
            1           8        11.8              9       13.2
            2          24        35.3             33       48.5
            4           2         2.9             35       51.5
            5          16        23.5             51       75.0
            6          17        25.0             68      100.0

                         Frequency Missing = 2
```

Bibliography

Air Force Center for Environmental Excellence (AFCEE). Peer Review Guidance Document. Draft 1.2. AFCEE/ERC, 29 Jan 96.

Air Force Materiel Command (AFMC). "1996 AFMC Progam Review." Slide Briefing. 24 August 1996.

American Consulting Engineers Council and American Society of Civil Engineers. Project Peer Review Guidelines. 1990.

American Consulting Engineers Council and American Society of Civil Engineers. Reviewer's Manual. 1990.

American Consulting Engineers Council, Coalition of American Structural Engineers. Technical Peer Review Manual. 1988.

American Society of Civil Engineers. Quality in the Constructed Project: A Guideline for Owners, Designers and Contructors, Volume 1. New York: ASCE, 1988.

Association of Soils and Foundations Engineers. A Peer Review Program Manual for Peer Reviewers and Participating Firms and Offices. Silver Spring: ASFE, 1990.

Atkinson, Malcolm. "Regulation of Science by 'Peer Review,'" Studies in History & Philosophy of Science, 25: 147-158 (April 1994).

Babbie, Earl. Survey Research Methods. Belmont: Wadsworth Publishing Company, 1990.

Barnes, Craig E., "Project Structural Peer Review in Massachusetts," Restructuring--America and beyond: Proceedings of Structures Congress XIII, Boston, MA, April 2-5, 1995: 1085-1088 (1995).

Beyer, Janice M., Roland G. Chanove, and William B. Fox. "The Review Process and the Fates of Manuscripts Submitted to AMJ," Academy of Management Journal, 38: 1219-1260 (1995).

Blank, Rebecca M. "The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence from the American Economic Review," American Economic Review, 81: 1041-1067 (December 1991).

Burnham, John C. "The Evolution of Editorial Peer Review," Journal of the American Medical Association, 263: 1323-1329 (9 March 1990).

Carlson, David A. "Special Inspections and Project Peer Review in Connecticut," Restructuring--America and beyond: Proceedings of Stuctures Congress XIII, Boston, MA, April 2-5, 1995: 1102-1105 (1995).

Chubin, Daryl E. and Edward Hackett. Peerless Science. New York: State University of New York Press, 1990.

Cicchetti, Domenic V. "Author's Response to Commentary on 'The reliability of peer review for manuscript and grant submissions: A cross-disciplinary investigation,'" Behavior and Brain Sciences, 16: 401-403 (1993).

Cicchetti, Domenic V. "The reliability of peer review for manuscript and grant submissions: A cross-disciplinary investigation," Behavior and Brain Sciences, 14: 119-186 (1991).

Cleary, J. D. "Blind versus nonblind review: survey of selected medical journals." Drug Intellective Clinical Pharmacy, 22: 601-602 (1988).

Cozzens, Susan E. "Program evaluation: Expert review in evaluating programs," Science and Public Policy: 71-81 (April 1987).

Creswell, John W. Research Design, Qualitative & Quantitative Approaches. Thousand Oaks: Sage Publications, Inc., 1994.

Cronbach, L. J. "Coefficient alpha and the internal structure tests," Psychometrika, 16: 297-334 (1951).

Daniel, H. D., Translated by William E. Russey. Guardians of Science: Fairness and Reliability of Peer Review. Weinheim, Germany: VCH, 1993.

Department of Energy. Procedures for Peer Review Assessments. DOE/ER-0491P. Washinton: U.S. Department of Energy, April 1991.

Department of the Air Force. Air Force Personnel Survey Program. AFI 36-2601. Washington: HQ USAF, 1 February 1996.

Department of the Air Force. Air University Sampling and Surveying Handbook: Guidelines for planning, organizing, and conducting surveys. Washington: HQ USAF, 1993.

Department of the Air Force. Environmental Information Management System. AFI 32-7002. Washington: HQ USAF, 31 May 1994.

Department of the Air Force. Environmental Quality. AFPD 32-70. Washington: HQ USAF, 20 July 1994.

Department of the Air Force. The Environmental Restoration Program. AFI 32-7020. Washington: HQ USAF, 19 May 1994.

Department of the Air Force. The Information Collections and Reports Management Program; Controlling Internal, Public, and Interagency Air Force Information Collections. AFI 37-124. Washington: HQ USAF, 29 March 1994.

Department of the Air Force. The Metrics Handbook. Washington: HQ USAF, August 1991.

Department of the Air Force. U.S. Air Force Installation Restoration Program Remedial Project Manager's Handbook. Washington: HQ USAF, December 1993.

Department of the Army: Program Evaluation Through a Peer Review Process. EP 1130-2-436. Washington: HQ USACE, 1 April 1994.

Devore, Jay L. Probability and Statistics for Engineering and the Sciences. 4th Ed. Belmont: Duxbury Press, 1995.

Dillman, Don A. Mail and Telephone Surveys: The Total Design Method. New York: John Wiley & Sons, 1978.

Doraiswami, M. and Andrew Towl. ASCI CASE COLLECTION. New York: ASIA Publishing House, 1963.

Feurer, Irene D., Gary J. Becker, Daniel Picus, Estella Ramirez, Michael D. Darcy, and Marshall E. Hicks. "Evaluating Peer Reviews," Journal of the American Medical Association, 272: 98-100 (13 July 1994).

Furlong, Robert. Headquarters Civil Engineering Environmental Restoration. Telephone interview. 2 November 1995.

Gallagher, Eugene B. "Editorial: Peer Review Always Helps But Never Cures," American Sociological Association, 30: iv-v (September 1989).

Garfunkel, Joseph M., Martin H Ulshen, Harvey J. Hamrick, and Edward E. Lawson. "Problems Identified by Secondary Review of Accepted Manuscripts." JAMA, 263: 1369-1371 (9 March 1990).

General Accounting Office. Peer Reveiw: EPA Needs Implementation Procedures and Additional Controls. Report Series GAO/RCED-94-89. Washington: Government Printing Office, February 1994.

Hatcher, Larry and Edward J. Stepanski. A Step-by-Step Approach to Using the SAS® System for Univariate and Multivariate Statistics. Cary: SAS® Institute Inc., 1994.

Hargens, Lowell L. "Variation in Journal Peer Review Systems," The Journal of the American Medical Association, 263: 1348 (9 March 1990).

Hirschman, David J. "Peer Review and the Contract Between Science and Society," The Environmental Professional, 16: 79-89 (1994).

Horrobin, David F. "The Philosophical Basis of Peer Review and the Suppression of Innovation," Journal of the American Medical Association, 263: (9 March 1990).

Johnson, Arthur W. "Standard Contracts for Project Peer Review and Special Inspection Services," Restructuring: America and Beyond: Preceedings of Structures Congress XIII, Boston, MA, April 2-5, 1995: 1097-1101 (1995).

Kachigan, Sam K. Multivariate Statistical Analysis. 2nd Ed. New York: Radius Press. 1991.

Knoll, Elizabeth. "The Communities of Scientists and Journal Peer Review," Journal of the American Medical Association, 263: 1330-1333 (9 March 1990).

Kronick, David A. "Peer Review in 18th-Century Scientific Journalism," Journal of the American Medical Association, 263: 1321 (9 March 1990).

Luce, R. Duncan. "Reliability is neither to be expected nor desired in peer review," Behavioral and Brain Sciences, 16: 399-400 (June 1993).

Marshall, Eliot. "Peer Review Reforms Get Good Review," Science, 265: 467 (22 July 1994).

McCabe, R.K. "Wanted: A Few Good Reviewers," Journal of Accountancy, 172: 109-114 (August 1991).

McNutt, Robert A., Arthur T. Evans, Robert H. Fletcher, and Suzanne W. Fletcher. "The Effects of Blinding on the Quality of Peer Review." JAMA, 263: 1371-1376 (9 March 1990).

Mundey, Jeff. Chief of Environmental Restoration Branch, HQ AFMC/CEVR. "FY96 Peer Review." Memorandum. Undated.

National Institute of Allergy and Infectious Diseases. Peer Review of Contract Proposals: Guidelines for Reviewers. National Institutes of Health, Department of Health and Human Services. July 1980.

Nesmith, Lynn. "GSA Establishes Peer Review For Major Public Projects," Architect Record, 181: 29 (October 1993).

Nunnally, P. E. Psychometric Theory. New York: McGraw-Hill, 1978.

"Official Releases: Peer review standards interpretations," Journal of Accountancy, 180: 118-119 (September 1995).

Oxman, Andrew D., Gordon H. Guyatt, Joel Singer, Charlie H. Goldsmith, Brian G. Hutchinson, Ruth A. Milner, and David L. Streiner. "Agreement Among Reviewers of Review Articles." Journal of Clinical Epidemiology, 44: 91-98 (1991).

Patton, Michael Q. Qualitative Evaluation and Research Methods. Newbury Park: The International Professional Publishers, 1990.

Peer Review Board of the American Institute of Certified Public Accountants. AICPA Peer Review Program Manual. Chicago: CCH Incorporated, 1995.

Preziosi, Dominic. "Reviewing Peer Review," Civil Engineering, 58: 46-48 (November 1988).

Rennie, Drummond and Annette Flanagin. "Peer Review in Prague: The International Congress on Biomedical Peer Review and Global Communications," Journal of the American Medical Association, 274: 986-987 (27 September 1995).

Rennie, Drummond and Annette Flanagin. "The Second International Congress on Peer Review in Biomedical Publication," Journal of the American Medical Association, 272: 91 (13 July 1994).

Schwartz, Samuel M., Donald W. Slater, Fred P Heydrick, and Gillian R. Woolett. "A report of the AIBS peer-review process for the US Army's 1994 breast cancer initiative," Bioscience, 45: 558-563 (September 1995).

Shrout, Patrick E. and Joseph L. Fleiss. "Intraclass Correlations: Uses in Assessing Rater Reliability," Psychological Bulletin, 86: 420-428 (1979).

Sinclair, J. D. “Drop censorship in science,” Behavioral and Brain Sciences, 16: 400-401 (June 1993).

Smith, Richard. “Promoting research into peer review,” British Medical Journal, 309: 143-144 (16 July 1994).

Stossel, T. P. “Reviewer Status and Review Quality,” The New England Journal of Medicine, 312 658-659 (7 March 1985).

Taubes, Gary. “Peer Review Goes Under the Microscope,” Science, 262: 25-26 (1 October 1993).

Tresch, Reneta. Qualitative research: Analysis types and software tools. New York: Falmer, 1990.

Weller, Ann C. “Editorial Peer Review in US Medical Journals,” Journal of the American Medical Association, 263: 1344-1347 (9 March 1990).

Zallen, Rubin M. and others. “Proposal for Stuctural Design Peer Review,” Journal of Performance of Constructed Facilities, 6: 186-204 (August 1992).

Zallen, Rubin M. “Proposal for Structural Design Peer Review,” Journal of Performance of Constructed Facilities, 4: 208-215 (November 1990).

Zuckerman, Harriet, and Robert K. Merton. “Patterns of Evaluation in Science: Institutionalization, Structure and Functions of the Referee System,” Minerva, 9: 66-100 (January 1971).

Vita

Captain Paul A. Schantz was born on 10 January 1967 in North Tarrytown, New York. He graduated from Liverpool High School in 1985 and entered undergraduate studies at Rensselaer Polytechnic Institute in Troy, New York. He graduated with a Bachelor of Science in Electrical Engineering in May 1989 and recieved a reserve commission in the United States Air Force.

Capt Schantz was first assigned to the 2851st Civil Engineering Squadron at Kelly Air Force Base, in San Antonio, Texas as a Design Engineer. While there he held both the Readiness Chief and Logistics Section Chief positions. In October 1992, Capt Schantz was transferred to the 11th Civil Engineering Operations Squadron at Elmendorf Air Force Base in Anchorage, Alaska.

In May 1995 he was awarded a Masters of Science in Engineering Management from the University of Alaska in Anchorage; he then entered the Graduate School of Engineering at the Air Force Institute of Technology. On 21 July 1995, Capt Schantz was certified as a registered Professional Engineer. On 23 February 1996, he recieved a regular commision. On 23 May 1996, he was recognized as a member of Tau Beta Pi, Engineering Honor Society.

Permanent Address: 4191 Orion Path
Liverpool, NY 13090

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** December 1996	**3. REPORT TYPE AND DATES COVERED** Master's Thesis

4. TITLE AND SUBTITLE Environmental Restoration Project Peer Review Process Assessment	**5. FUNDING NUMBERS**
6. AUTHOR(S) Paul A. Schantz, Captain, USAF	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology 2750 P Street WPAFB, OH 45433-7765	**8. PERFORMING ORGANIZATION REPORT NUMBER** AFIT/GEE/ENV/96D-18
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Maj Dan Welch AFCEE/ERC 8001 Arnold Drive Brooks AFB, TX 78235-5357	**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited	**12b. DISTRIBUTION CODE**

13. ABSTRACT *(Maximum 200 words)*

Installation Restoration Program (IRP) projects often cost over $250 thousand, and projects over $1.5 million are common. To ensure these projects are risk based, technically sound, and cost effective, the Air Force instituted a peer review program in 1992. The objective of this research is to describe and analyze the peer review process.

Through triangulation of data from interviews, observations, official and academic documents, and surveys, seven constructs were discovered: focus, agenda, facilitator, written preparation, oral presentation, team characteristics, and reviewer characteristics.

A questionnaire gathered perceptions of peer review effectiveness--the criterion variable--and of the seven constructs. There were a total of 141 surveys administered with a 50% response rate (N = 70).

Linear regression is next used to assess the predictiveness of the seven constructs. Oral preparation and focus predict 68% of the variance in peer review effectiveness; no other constructs were significant. Variation in the peer review process between major commands is explored.

Last, recommendations are made to improve the current peer review process.

14. SUBJECT TERMS Peer Review, Linear Regression, Survey, Questionnaire, Installation Restoration Program, Environmental, Project Management, Environmental Management, Environmental Metrics	**15. NUMBER OF PAGES** 216
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UL

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet ***optical scanning requirements***.

Block 1. Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C -	Contract	**PR** -	Project
G -	Grant	**TA** -	Task
PE -	Program Element	**WU** -	Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. ***(If known)***

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD -	See DoDD 5230.24, "Distribution Statements on Technical Documents."
DOE -	See authorities.
NASA -	See Handbook NHB 2200.2.
NTIS -	Leave blank.

Block 12b. Distribution Code.

DOD -	Leave blank.
DOE -	Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
NASA -	Leave blank.
NTIS -	Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

★U.S.GPO:1993-0-336-043